





UNIVERSITY OF ILLINOIS AT  
CHICAGO  
801 SO. MORGAN  
CHICAGO, IL 60607

CHICAGO CIRCLE



LIBRARY

This book is the gift of

Dreyfus Fund

UNIVERSITY of ILLINOIS







Q

1

A773

V. 3

1950

PER



AUSTRALIAN JOURNAL  
OF  
SCIENTIFIC RESEARCH

SERIES A  
PHYSICAL SCIENCES

VOLUME 3

MELBOURNE  
1950



# AUSTRALIAN JOURNAL OF SCIENTIFIC RESEARCH

Published by the Commonwealth Scientific and Industrial Research Organization and the Australian National Research Council

Issued in two Series :

Series A – Physical Sciences

(Four issues each year appearing March, June, September, December)

Series B – Biological Sciences

(Four issues each year appearing February, May, August, November)

Price : Each series 30/- per annum, separate issues 7/6 each

---

## EDITORIAL BOARD

Chairman and Editor : N. S. Noble

Members : Professor F. M. Burnet  
Professor W. J. Dakin  
Professor E. J. Hartung  
Professor L. H. Martin  
Professor J. G. Wood

---

All enquiries and manuscripts should be forwarded to :

The Editor,  
Australian Journal of Scientific Research,  
Commonwealth Scientific and Industrial Research Organization,  
314 Albert Street, East Melbourne, C.2, Victoria.



## CONTENTS

NUMBER 1, MARCH 1950

	PAGE
On the Dynamics of the General Atmospheric Circulation. By C. H. B. Priestley .. .. .	1
Galactic Radiation at Radio Frequencies. I. 100 Mc/s. Survey. By J. G. Bolton and K. C. Westfold .. .. .	19
Radio-Frequency Radiation from the Quiet Sun. By S. F. Smerd ..	34
Solar Radiation at a Wavelength of 3.18 Centimetres. By H. C. Minnett and N. R. Labrum .. .. .	60
Deformation and Recrystallization of Alloys containing Two Phases. By L. M. Clarebrough .. .. .	72
Self-Diffusion in Tin Crystals. By P. J. Fensham .. .. .	91
Self-Diffusion in Polycrystalline Tin. By P. J. Fensham . . .	105
Conjugated Compounds. II. Simple Potential Energy Functions, Absorption Spectra, and Ionization in Linear Polyenes. By N. S. Bayliss .. .. .	109
A Colorimetric Method for the Determination of Esters. By Adrienne R. Thompson .. .. .	128
Studies of the Wallach Transformation. I. The Products of the Reaction. By P. H. Gore and G. K. Hughes .. .. .	136
Alkaloids of the Australian Rutaceae. The Structure and Reactions of Acronycidine. By F. N. Lahey, J. A. Lamberton, and J. R. Price..	155
Corrigenda .. .. .	172



## NUMBER 2, JUNE 1950

	PAGE
Estimation of the Error at a Point on a Least-Squares Curve. By P. G. Guest .. .. .	173
Remarks on the Latitude Effect of Cosmic Rays at Sea-Level. By H. D. Rathgeber .. .. .	183
The Formation of Rain by Coalescence. By E. G. Bowen .. ..	193
Observation of Precipitation with an Airborne Radar. By E. J. Smith..	214
Survey of Galactic Radio-Noise at 200 Mc/s. By C. W. Allen and C. S. Gum.. .. .	224
Galactic Radiation at Radio Frequencies. II. The Discrete Sources. By G. J. Stanley and O. B. Slee .. .. .	234
Galactic Radiation at Radio Frequencies. III. Galactic Structure. By J. G. Bolton and K. C. Westfold .. .. .	251
The Polarization of Thermal "Solar Noise" and a Determination of the Sun's General Magnetic Field. By S. F. Smerd .. .. .	265
Radiative Heat Transfer in the Air near the Ground. By E. L. Deacon..	274
The Superstructure in the $\alpha$ Phase of Silver-Magnesium Alloys. By L. M. Clarebrough and J. F. Nicholas .. .. .	284
The Catalytic Dehydration of 2,3-Butanediol to Butadiene. II. Adsorption Equilibria. By M. E. Winfield .. .. .	290
Liquid-Vapour Equilibria. I. The Systems Carbon Tetrachloride- <i>cyclo</i> - Hexane and Water-Acetic Acid. By I. Brown and A. H. Ewald	306
The Synthesis of Polymers in Reduced Wool. By M. Lipson and R. J. Hope .. .. .	324
Some Bridged Derivatives of 4-Piperidone. By E. F. L. J. Anet, G. K. Hughes, Diana Marmion, and E. Ritchie .. .. .	330
A Synthesis of <i>iso</i> Pelletierine and Methyl <i>iso</i> pelletierine. By E. F. L. J. Anet, G. K. Hughes, and E. Ritchie .. .. .	336
The Chemical Constituents of Australian <i>Xanthoxylum</i> Species. I. Suberosin, 6-( $\gamma\gamma$ -Dimethylallyl)-7-Methoxycoumarin. By Jean Ewing, G. K. Hughes, and E. Ritchie .. .. .	342
The Alkaloids of <i>Pleogyne cunninghamii</i> . By F. A. L. Anet, G. K. Hughes, and E. Ritchie .. .. .	346
A Correction .. .. .	350



## NUMBER 3, SEPTEMBER 1950

	PAGE
Experimental Designs Balanced for Pairs of Residual Effects. By E. J. Williams .. .. .	351
Estimation of the Errors of the Least-Squares Polynomial Coefficients. By P. G. Guest .. .. .	364
Equivalent Path and Absorption for Electromagnetic Radiation in the Solar Corona. By J. C. Jaeger and K. C. Westfold .. .. .	376
Observations of the Spectrum of High-Intensity Solar Radiation at Metre Wavelengths. I. The Apparatus and Spectral Types of Solar Burst Observed. By J. P. Wild and L. L. McCreedy .. .. .	387
Observations of the Spectrum of High-Intensity Solar Radiation at Metre Wavelengths. II. Outbursts. By J. P. Wild .. .. .	399
Atmospheric Noise Levels at Radio Frequencies near Darwin, Australia. By D. E. Yabsley .. .. .	409
The Pressure Dependence of the Thermal Conductivity of Polyatomic Gases at 0° C. By W. G. Kannuliuk and Heather B. Donald .. .. .	417
Asymptotic Expressions for the Energies of Certain Long Molecules. By R. D. Brown .. .. .	428
A Study of the Oxidation of Ethylene to Ethylene Oxide on a Silver Catalyst. By K. E. Murray .. .. .	433
The Reaction between Acetone and Ammonia. II. Isomeric Oximino-ketones related to Diacetoamine. By N. C. Hancox .. .. .	450
Sulphonamides. I. "Marfanil" and its <i>o</i> - and <i>m</i> -Isomers. By S. J. Angyal and S. R. Jenkin .. .. .	461
Dunnione and Related Naphthoquinones. I. The <i>iso</i> Dunnione Series. By R. G. Cooke and T. C. Somers .. .. .	466
Dunnione and Related Naphthoquinones. II. Synthesis of <i>iso</i> Dunniol and <i>dl</i> -Dunnione. By R. G. Cooke .. .. .	481
Dunnione and Related Naphthoquinones. III. Rearrangement of Hydroxynaphthoquinones to Indenone Carboxylic Acids. <i>allo</i> -Dunnione. By R. G. Cooke and T. C. Somers .. .. .	487
The Synthesis of Melicopicine and Some Trimethoxy-10-Methylacridones. By G. K. Hughes, K. G. Neill, and E. Ritchie .. .. .	497
The Existence of a Uronic Acid Ester in Young Wood of <i>Eucalyptus regnans</i> . By D. H. Foster, Gerda Schwerin, and W. E. Cohen .. .. .	504
<i>Short Communications</i>	
Nuclear Emulsion Technique. By V. D. Hopper .. .. .	512
A New Source of "l-Quercitol" (Viburnitol). By Jean Ewing, G. K. Hughes, and E. Ritchie .. .. .	514
The Chemistry of Western Australian Plants. III. Ursolic Acid from <i>Anthotroche</i> Species. By W. Bottomley and D. E. White .. .. .	516



## NUMBER 4, DECEMBER 1950

	PAGE
Interactions between Neutron and Proton. By P. Swan .. ..	519
Geometrical Optics of Concave Mirrors and of Combinations of Mirrors. By J. Shearer .. .. .	532
Observations of the Spectrum of High-Intensity Solar Radiation at Metre Wavelengths. III. Isolated Bursts. By J. P. Wild ..	541
Polarography with Alternating Currents. I. Outline of Theory, Apparatus, and Technique. By B. Breyer, F. Gutman, and S. Hacobian .. .. .	558
Polarography with Alternating Currents. II. A.C. Polarography of Cadmium, Zinc, Lead, Thallium, Indium, Bismuth, and Antimony. By B. Breyer, F. Gutman, and S. Hacobian .. .. .	567
The Crystal Structure of Indium Monobromide. By N. C. Stephenson and D. P. Mellor .. .. .	581
The Iodination of Aromatic Compounds. IV. The Iodination of Aromatic Hydrocarbons and Nuclear-substituted Phenols. By L. Jurd ..	587
The Ultraviolet Absorption Spectra of the Acridone Alkaloids. I. Compounds containing the Acridone Nucleus. By R. D. Brown and F. N. Lahey .. .. .	593
The Ultraviolet Absorption Spectra of the Acridone Alkaloids. II. Compounds related to 4-Quinolone. By R. D. Brown and F. N. Lahey .. .. .	615
Colouring Matters of Australian Plants. I. The Structure of Droserone. By R. G. Cooke and W. Segal.. .. .	628
A Synthesis of Sparteine and Some Related Substances. By E. F. L. J. Anet, G. K. Hughes, and E. Ritchie .. .. .	635
Aromatic Aldehydes from the Oxidation of Some Australian Woods and their Chromatographic Separation. By D. E. Bland, G. Ho, and W. E. Cohen .. .. .	642
Index to Volume 3 .. .. .	649









# ON THE DYNAMICS OF THE GENERAL ATMOSPHERIC CIRCULATION

By C. H. B. PRIESTLEY\*

[*Manuscript received January 19, 1950*]

## *Summary*

The annual flow of angular momentum across latitudes 30–35° effected by north and south currents lying side by side (cyclonic and anticyclonic “eddy”) appears insufficient to balance the angular momentum generated at the surface in lower latitudes. The discrepancy suggests that there must be also a mean poleward drift of air at upper tropospheric and lower stratospheric levels across the latitudes of the mean surface high pressure belt, with a counterdrift which occurs at least in part in near-surface levels. A variety of auxiliary evidence supports the existence of these flows, which are indicated as intermittent or fluctuating components with mean speed over a long period of the order of a few miles per hour and so, at upper levels, only on the margin of direct observation. The origin of the poleward component is associated with the strong anticyclonic shear to equatorward of the jet stream axis.

Since these drifts are found in latitudes where the vertical wind shear is greatest, they constitute an important agency in the transport of angular momentum from its source region in the tropics to the sink region of higher latitudes. This significant contribution is masked in studies of interchange which invoke the assumption of geostrophic wind.

The “index” representing the total strength of the tropical easterly surface winds, which fluctuates considerably, relates closely to the rate of generation of angular momentum in the tropics. On the basis of the drifts referred to above, a descriptive theory of the index cycle is developed, which links the fluctuations of index with the fluctuations in strength and in latitude of the jet stream. Quantitative checks are applied to this theory wherever possible and no discrepancy appears, while the normal sequence of synoptic events agrees broadly with that ascribed to the mutual interactions of surface wind strength, jet-stream, and drift.

## I. INTRODUCTION

The frictional stress  $\tau$  exerted by the earth's surface on the atmosphere is opposed in direction to the surface wind. If the wind speed is  $V$ , then (see, e.g. 1)

$$\tau = k\rho V^2, \dots\dots\dots (1)$$

where  $\rho$  is air density and  $k$  a number which depends on the nature of the surface and the height at which  $V$  is measured. Let  $\bar{\tau}_x$  denote the annual mean eastward component of  $\tau$ . Knowing the order of magnitude of  $k$  and the surface wind distribution over the globe,  $\bar{\tau}_x$  can be assessed as a function of latitude  $\phi$ . The values found in certain zones are such as could destroy the circulation of those zones in a matter of days(2), yet there is no immediately identifiable trend in the

\* Section of Meteorological Physics, C.S.I.R.O., Melbourne.

pattern of general circulation. This relative constancy of pattern in the face of such strong torques finds expression in the equation of moments

$$\int_{-\pi/2}^{+\pi/2} \bar{\tau}_x \cos^2 \varphi \, d\varphi = 0 \quad \dots \dots \dots (2)$$

A similar expression must be closely satisfied for mean values over periods much shorter than a year, though not necessarily for instantaneous values.

Since the surface zonal winds in polar regions are not disproportionately strong, it follows from (2) that the stresses in these regions are of little moment. The problem of the maintenance of the mean zonal circulation is therefore concerned primarily with the maintenance of the westerly and easterly zones of surface winds in middle and low latitudes respectively. It will be assumed that, to a first approximation, the circulation is planetary, i.e. that the frequency distribution of winds, in a given latitude, is independent of longitude and is similar in the two hemispheres. The importance of geographical factors can only be resolved by more detailed and specific local studies than are here to be attempted.

Applying the principle of balance to zones extending from the equator to  $\pm\varphi_0$

$$2\pi \int_{-\varphi_0}^{\varphi_0} = 2\pi \int_0^{\varphi_0} R^3 \bar{\tau}_x \cos^2 \varphi \, d\varphi = \text{mean poleward flux of angular momentum across latitude } \varphi_0 \text{ or } -\varphi_0, \dots (3)$$

$R$  being the mean radius of the earth. This integrated equation is independent of any assumption as to conservation of angular momentum of the individual moving elements (see, e.g. 3).

It was argued by Jeffreys that the main agents of flux are the rotations of cyclonic dimensions in horizontal planes, at least in the latitudes of the prevailing surface westerlies. Extended to all latitudes, this view encounters the difficulty that the maximal flux must occur across the high pressure belt around  $30-35^\circ$  latitude, which is the belt of minimal surface disturbance; nevertheless it is widely held that this is effected by a marked and systematic asymmetry in the individual anticyclones. But there is no published quantitative evidence to show that horizontal rotations constitute the sole or even the major agency of momentum exchange across these latitudes. The possible importance of other agencies must still be considered, particularly since the short period *variations* in flux are of great dynamical significance.

The classical theory that zonal motions are maintained largely by circulations in meridional planes does not enjoy current favour. In fact the bulk of recent and current work in this field disregards the possibility of any such contribution, studies of momentum exchange being based on the assumption of geostrophic wind. The present work is based almost entirely on actual wind data, in the hope thereby of reaching a proper perspective on the classical theory and also on the validity of the geostrophic approximation in this context. The results suggest most strongly that, at least in the critical latitudes of the high pressure belt, a significant poleward flux of angular momentum is effected by a slight



mean poleward drift of air at some levels accompanied by an equivalent mass drift equatorward at others, i.e. by the components in meridional planes of a mean toroidal circulation. In some of the arguments and tests here advanced, the *whole* of the flux will be assigned to the toroidal component. This is an extreme hypothesis designed to emphasize the probable importance of the toroidal component and set an upper limit to its magnitude. This is to say, one extreme hypothesis is at times invoked in order to call attention to the limitations of the other. By so doing it is not meant to imply that agents other than meridional circulations are in practice negligible; a fuller understanding of the complex processes in operation is to be found between the two extremes.

## II. MAGNITUDES AND TERMINOLOGY

The first step requires the assessment of the maximal value of angular momentum flux from (1) and (3) with  $\varphi_0 \approx 30^\circ$ . Surface wind frequencies over the oceans covering many years are available.\* The annual mean easterly component and mean speed of resultant surface winds were determined for  $5^\circ$  latitude zones from  $0$  to  $30^\circ$  S., and their products obtained as in Table 1. The values of  $\bar{\tau}_x$  so derived are not truly representative, since they are based on  $\bar{V}^2$  and not on  $\bar{V}$ , but in these regions of steady wind no great error is involved on this account.

TABLE 1  
ANNUAL MEAN SURFACE WIND DATA

Latitude .. ..	0-5	5-10	10-15	15-20	20-25	25-30 °S.
Easterly component ..	170	260	450	435	350	180 cm./sec.
Resultant speed ..	450	440	630	590	590	610 cm./sec.
Product .. ..	0.75	1.15	2.8	2.6	2.1	$1.1 \times 10^6$ cm. <sup>2</sup> /sec. <sup>2</sup>

Evaluation of  $k$  is less straightforward (1, p. 260).  $k$  can be related with "roughness height"  $z_0$  and the height  $z$  at which wind is measured by

$$k = \frac{0.16}{\left( \log_e \frac{z}{z_0} \right)^2}.$$

Baltic Sea data(4) yield  $z_0 = 0.06$  cm. while Rossby and Montgomery(5) give  $z_0 = 0.6$  cm. over sea. The height to which  $z$  refers is also uncertain, but putting  $z = 700$  cm. as supposedly representative we find  $k = 0.0018$  or  $0.0032$  according to the value of  $z_0$  adopted, and the uncertainty of  $z$  does not greatly affect this range. We shall adopt the mean value  $0.0025$ , whence the poleward flow of angular momentum across latitude  $30^\circ$  S. in one year is equal to

$$+1.3 \times 10^{34} \text{ g. cm.}^2/\text{sec.} \dots\dots\dots (4)$$

A considerable margin of error is to be recognized in this estimate. The increase in friction over some land areas and very rough seas has not been allowed for, so that (4) may be considerably too small; but only if it were too *large* by

\* British Admiralty Publ. Item No. 163, corrected up to 1943.

about a threefold factor, which scarcely seems possible, would the arguments which follow be seriously weakened. Starr\* (unpublished data) has recently reported a generation between 10 and 30 °N. for a period of only two months (Dec. 1945 and Jan. 1946) of amount  $3.3 \times 10^{33}$ , estimated by quite a different technique of assessment. A further contribution would arise for the zone 0 to 10° N. while the generation by pressure differences across mountain barriers, not considered here, provides a further reinforcement. On this basis the annual generation should be, if anything, in excess of (4).

Let now  $\bar{u}$ ,  $\bar{v}$ ,  $\bar{\Omega}$  denote respectively the mean annual westerly and poleward wind components and angular momentum per unit mass. The averages are taken at fixed pressure levels and will be functions of pressure  $p$  and of  $\varphi$ .  $p_0$  will denote surface pressure,  $\omega$  the angular velocity of the earth's rotation,  $g$  gravity acceleration, and  $z$  height. Neglecting quantities of order  $z/R$ ,

$$\bar{\Omega} = R^2 \omega \cos^2 \varphi + R \bar{u} \cos \varphi \dots\dots\dots (5)$$

That part of the flux ascribable to the mean components is

$$\frac{2\pi R \cos \varphi}{g} \int_0^{p_0} \bar{v} \bar{\Omega} dp = \frac{2\pi R^2 \cos^2 \varphi}{g} \int_0^{p_0} \bar{v} \bar{u} dp, \dots\dots\dots (6)$$

so that potentially the most significant levels are those where  $\bar{u}$  is greatest. Equation (6) does not apply for any limits of integration but only when the whole depth of atmosphere is considered, for its derivation depends on the constancy of the mean pressure distribution through the condition

$$\int_0^{p_0} \bar{v} dp = 0 \dots\dots\dots (7)$$

Mean annual westerly winds show a maximum around latitude 30° at a height of 12–13 km. (6, 7), with mean winter speed exceeding 100 knots. In order to estimate roughly what magnitude of "drift"  $\bar{v}$  is required, it will be assumed constant from 400–100 mb. throughout which  $\bar{u} = 50$  knots and the counter drift, assumed to take place through layers of much smaller  $\bar{u}$ , will be neglected in the first approximation. From (4) and (6), then, it can be evaluated that

$$\bar{v} \approx +300 \text{ cm./sec.} \dots\dots\dots (8)$$

as a rough mean value for the whole layer, on the extreme hypothesis referred to in the previous section.

On the geostrophic assumption  $\bar{v} = 0$  automatically at all levels. But it is evident from the above that a  $\bar{v}$  of only two or three miles per hour would play an important part in the dynamics of the general circulation, and the possibility of such a  $\bar{v}$  cannot be discounted. Attention must therefore be given to its potential role in relation to the two major features of the mean pattern, namely :

(i) The belts of surface easterly winds in low and westerlies in middle latitudes.

\* Investigation of the general circulation of the atmosphere, M.I.T. Final Report, Aug. 1949.



(ii) Upper tropospheric and lower stratospheric westerly winds which attain maximum speed in a zone whose location in latitude varies with time. In view of the large magnitude of these winds relative to  $\bar{v}$ , it is only to be expected that the "drift" should so far have escaped direct observation and, for the same reason, its existence can be established only by means of a concordant body of indirect evidence.

It is convenient here to define certain terms which will be frequently employed. A belt of latitude girdling the earth is termed a *zone*, the part of a zone between two heights or pressure levels a *layer*. The main zones of (i) will be described as *easterly* and *westerly* or, when these terms may be confusing, as *tropical* and *temperate*, words which here do not imply any fixation in latitude. In conformity with current practice we shall refer to occasions when the belts in (i) are intense or weak as occasions of *high index* and *low index* respectively,\* to the phenomenon of (ii) as the *jet stream*, and to the region of highest instantaneous wind as the *axis* of the jet. We shall continue to refer to  $\bar{v}$  as the *drift*, applying the concept specifically to the latitude of the high pressure belt. The convention is that a poleward drift be positive, the word *counter-drift* being employed in the opposite sense.

### III. DYNAMICAL THEORY OF THE INDEX FLUCTUATIONS

On the basis of the drift a theory can be constructed capable of accounting for the broad pattern of index fluctuations and of explaining the manner in which these fluctuations interact with well-known jet stream phenomena. This theory will now be outlined, critical discussion and quantitative tests being postponed to later sections of the paper.

As such a theory must deal with instantaneous rather than annual mean values, the bars over the symbols will be dropped, the symbols still however referring to the instantaneous space-mean properties of a zone. The success of the theory then provides a measure of the extent to which an axially symmetric model of the circulation is legitimate and self-sufficient.

The high pressure belt will be identified by  $\varphi = \varphi_0$ . From the observed large fluctuations of the index(8) it appears that the flux of angular momentum across latitude  $\varphi_0$  is at most times appreciably out of adjustment with the rate of generation by surface friction in the easterly zone. Theoretical work(9) indicates that conditions of instability leading to quasi-horizontal rotations are almost always present whereas the type of instability leading to drift arises less frequently, and it seems natural to associate the spasmodic behaviour of the index mainly with the latter phenomenon especially since the periods of the fluctuations can be shown to be of the same order of magnitude as the time taken to build up the instability. While recognizing that other processes may be contributory, it is reasonable to test this supposition by constructing a theory on the basis of the extreme hypothesis of Section I. The trend of the circulation will depend on whether the efflux of angular momentum is temporarily less or

\* It has been shown(8) that there is a strong positive correlation between the strength of the easterlies in low and that of the westerlies in middle latitudes.

greater than the rate of generation, and we define states of *weak*, *balanced*, and *strong drift* according as

$$\cos \varphi_0 \int_0^{\varphi_0} (v\Omega)_{\varphi_0} d\varphi \stackrel{<}{\stackrel{>}{\approx}} gR^2 \int_0^{\varphi_0} \tau_x \cos^2 \varphi d\varphi \dots \quad (9)$$

respectively.

The criterion for instability of the type invoked is (10)

$$\partial\Omega/\partial\varphi > 0 \dots\dots\dots (10)$$

The differentiation is, strictly, along an isentropic surface, but applies approximately to horizontal or isobaric differentiation, the approximations becoming exact at the level of maximum wind. When the critical shear is exceeded as in (10) the orbits of disturbed elements become hyperbolic and there results a net horizontal outflow in both directions from the unstable zone. If the instability be the main cause of the mean drift, the poleward sense of the drift requires that instability should occur more frequently to equatorward of  $\varphi_0$ . The form of (10), written as

$$\partial u/\partial\varphi > 2\omega R \sin \varphi + u \tan \varphi$$

does indicate an *a priori* probability that instability should occur more frequently in the lower latitudes. But the mechanism so far described is not complete, for a sustained drift would not by itself lead to a condition in which (10) is satisfied, and the origin of the instability remains to be described.

Consider, first, a time when the zonal index is high for which, from (9), there is an *a priori* probability of weak drift which we shall assume does in fact hold. It follows that the air in low latitudes will on the whole gain westerly momentum, while the circulation in middle latitudes will tend to decline. The manner in which the momentum generated at the surface is shared by different layers will play an important part in the determination of subsequent events, but for the present it is sufficient to recognize that the surface-generated momentum is so shared. That part which is absorbed in the surface layers operates, in both zones, to reduce the index: that part which is absorbed in higher layers brings about an equatorward shift in the jet stream axis. So long as the axis is in high middle latitudes it is far removed from any positive source of angular momentum. As it moves to lower latitudes, however, the jet is able to draw increasingly on the positive angular momentum generated at the surface in the tropical zone. Movement of the jet to low latitudes is therefore accompanied by an increase in speed and in the anticyclonic shear on the equatorward side, the zonal speed in very low latitudes being restricted by the strong frictional coupling with the earth's surface near the equator (11).

Thus the sense of the angular momentum generated at the ground provides one physical influence which strengthens the *a priori* probability that the instability will occur preferentially in the tropical zone. The net drift will then be positive.

The onset of instability on the equatorward side of the high pressure belt temporarily restores the drift. The index will by this time be low, and the conditions those defined by (9) as strong drift. Westerly momentum will be transferred, faster than it is generated, from upper levels of the tropical to



upper levels of the temperate zone, a process signalled by a poleward displacement of the jet stream axis. This net loss and gain of angular momentum by the tropical and temperate zones respectively must subsequently make their influence felt at the surface in such a sense as would effect a recovery of the index. If, as will be established later, the return-drift implied by (7) takes place at least in part near the surface, this will aid further in restoring the index.

There results a cyclic index oscillation of which both the building and the declining phase will, in practice, occupy a finite time owing to the lags and necessary adjustments in various parts of the mechanism. On the cessation of instability the drift will fall off and eventually, as defined by (9), will become weak, while the revival of the index completes the cycle of processes.

Tests of the theory above outlined, which constitute the remainder of this paper, are of three types: quantitative checks on the internal consistency of the theory especially at points of conflict with others previously advanced (Sections IV and V), examination of the evidence for the drift (Section VI), and comparison of the pictured sequence of events with that actually observed (Section VII). Considerable detail is put forward but finality is not claimed, for the theory can be no surer than the foundations on which it is built and these display two main weaknesses. The concept of dynamical instability applies only to particle displacement and must eventually be extended to allow for the pressure modification resulting from the displacement of a finite mass of air; this may overcome the difficulty that the inertial ellipse appears to prohibit any great lateral penetration of a dynamically stable zone. Lack of quantitative appeal to thermodynamic influences constitutes the second basic weakness of this as of other mainly dynamical theories. These influences may in themselves tend to bring about a poleward drift at high levels (see, e.g. 12), so that the present theory contributes principally to an understanding of the allied *fluctuations* in drift and in the more readily identified features of the circulation pattern.

#### IV. THE PROCESS OF THERMAL ADJUSTMENT

The theory here advanced diverges in important respects from that put forward by Jeffreys(13) which has long held wide acceptance. The first point of conflict relates to the manner in which the angular momentum generated at the surface is redistributed through the overlying layers. Jeffreys regarded the temperature distribution as predetermined, and the redistribution of angular momentum thereby restricted to a form which would leave the vertical variation of wind unaltered. Such a postulate is only consistent with the geostrophic assumption, and in practice it must be recognized that the temperature and wind fields are determined by mutual adjustments.

The redistribution of  $\Omega$  is accordingly to be regarded as a process both of dynamical and thermal readjustment, the thermal contrast and vertical wind shear being subject with some possible lag in phase to the degree of exchange across the high pressure belt. To follow the consequences of this viewpoint we revert to the present theory at the first phase of the cycle, a period of weak drift commencing with an epoch of high index. It is known(6) that in periods of low index the anticyclonic belt is much broken by strong deep meridional

currents (horizontal rotations) ; whereas so long as the index remains high this mechanism of heat interchange between the two main zones is suppressed, and we may expect a progressive strengthening of the thermal contrast between latitudes. This lends support to the present theory. The process is perhaps best displayed schematically, using arrows pointing left or right to represent respectively easterly and westerly motion and accelerating influences as under.

			Wind	Dynamic Influence (Acceleration)	Thermal Influence (Relative Acceleration)
Tropical Zone :					
High Level	..	..	—→	—→	—→
Surface	..	..	←—	—→	←—
Temperate Zone :					
High Level	..	..	—→	←—	—→
Surface	..	..	—→	←—	←—

The dynamic influence is taken for all layers to be of the same sense, determined by the direction of the surface wind in the zone. The thermal influence is purely relative as between one layer and another. Although for this reason the arrows in the final column lack absolute significance, we are led to regard the upper layer in the tropics, where the westerly wind will increase, and the surface layer in middle latitudes, where the index will decline, as the more " active " regions in the process.

It is further suggested that this process of dynamic-thermal adjustment provides the explanation for the greater amplitude and regularity of index cycles in winter as compared with summer conditions(8). During periods when thermal interchange across latitude  $30^\circ$  is suppressed, it may be seen from Angot's insolation table quoted by Brunt(1, p. 112) that thermal gradients across this latitude belt will build up far more rapidly in winter than in summer.

The same principle holds for the readjustment in the second phase of the cycle, which commences with the lateral release of momentum to higher latitudes at an epoch of low index. With the promotion of interchange between the two zones, thermal contrasts are reduced to a value below that capable of supporting the vertical wind shears which have accumulated and which must now decline. Dynamic influences are also reversed since the tropical zone now experiences a net loss and the temperate zone a net access of angular momentum. Thus, as contrasted with the scheme above, the arrows in both the acceleration columns are reversed. The active regions remain as before, momentum is lost from upper levels in low latitudes and the surface circulation in the westerly zone is speeded up so as to revive the index.

As a quantitative point of appeal we require some estimate of the duration of the cycle: we select the time  $t$  taken for the development of upper-level instability from the mean state, which has been documented for regions of the Northern Hemisphere(6, 14). Extracting the mean winter zonal winds at the strongest level (12-13 km.) we denote by  $\Delta u$ ,  $\Delta \Omega$  the increments required to bring these winds to a condition of constant  $\Omega$ , where

$$\Delta \Omega = R \cos \varphi \cdot \Delta u.$$



Both sources of data give zero wind at approximately  $10^\circ$  latitude, indicating the boundary of the equatorial zone of strong coupling with the surface, and the constant  $\Omega$  profile is accordingly based on  $u=0$  at  $\varphi=10^\circ$ . Evaluation of  $\Delta u$  from  $10$  to  $25^\circ$  latitude is given in Table 2. Assuming that these increments

TABLE 2  
VELOCITY INCREMENTS  $\Delta u$ =WIND FOR CONSTANT  $\Omega$  MINUS MEAN WIND

Latitude	..	..	$10^\circ$	$15^\circ$	$20^\circ$	$25^\circ$	N.
$u$ (Willett)	..	..	0	17	35	50	m./sec.
$u$ (Hess)	..	..	0	(13)	25	30	"
Average	..	..	0	(15)	30	40	"
$u$ (constant $\Omega$ )	..	..	0	18	43	76	"
$\Delta u$	..	..	0	3	13	36	"

( ) denotes interpolation.

are common to the same layer as used in deriving (8), the total increment in angular momentum is

$$\begin{aligned} & \frac{2\pi R^3}{g} \int_{\varphi=10^\circ}^{\varphi=25^\circ} \int_{p=100 \text{ mb.}}^{p=400 \text{ mb.}} \cos^2 \varphi \cdot \Delta u \cdot dp \cdot d\varphi \\ & \approx 5 \times 10^{29} \int_{10^\circ}^{25^\circ} \cos^2 \varphi \cdot \Delta u \cdot d\varphi, \\ & \approx 1.2 \times 10^{32} \text{ g. cm.}^2/\text{sec.} \end{aligned}$$

Dividing this by (4) we obtain

$$t \approx \frac{1.2 \times 10^{32}}{1.3 \times 10^{34}} \text{ years} \approx 3 \text{ to } 4 \text{ days.}$$

This represents a crude estimate of the quarter-period of the cycle of events as here envisaged. In view of large tolerances, significance can be claimed only for the order of magnitude of this estimate, but the same may be said of the comparable estimate(13) that the surface circulation, if not reinforced, would be destroyed by friction in about 12 days. It might be added that on theoretical grounds the annihilation of the index would be asymptotic in time whereas the approach to  $\partial\Omega/\partial\varphi=0$  would not be so. Such considerations, though inevitably approximate, show that instability may often or even normally supervene to prevent destruction of the index, and a satisfactory quantitative basis for the new theory is established.

## V. FLUX EVALUATIONS

Equation (7), which represents the mean mass balance between latitudes, implies that the necessary transfer of angular momentum must be effected by small or large scale *rotations* of air. In Jeffreys's work and more recent studies based on the geostrophic assumption the whole of this transfer is assigned to the juxtaposition of horizontal eddies on the cyclonic or larger (long-wave) scale. In the present theory at least a significant part of the flux across  $\varphi_0$  is attributed

to the drift and counter-drift, i.e. to mean rotations in meridional planes. Since the flux of momentum, or of angular momentum, effected by eddies in horizontal planes can be directly determined from actual wind data(15), this issue is also one on which a quantitative appeal is possible.

It is convenient in evaluation to consider separately the effect of transient and transitory horizontal rotations (*eddy flux*) and of the seasonal or semi-permanent horizontal features (*standing-eddy flux*), as the former must be obtained from regular daily wind data at individual stations and the latter from world-wide seasonal charts.

The value of the eddy flux of momentum at a given place can be represented as an integrated Reynolds stress in the form

$$\frac{1}{g} \int_0^{p_0} \overline{v'u'} dp,$$

the data for a year being divided into six periods of two months, and the dash denoting daily departures from the mean value of the period. For conversion into the units employed in (4) we note that a stress of 1 dyne/cm. is equivalent to a flow of angular momentum of amount

$$8 \times 10^{25} \cos^2 \varphi \text{ g. cm.}^2/\text{sec.}$$

in one year. The calculation is laborious and suitable data, with accurately measured winds extending up to and beyond the tropopause, not plentiful. Full calculations have been carried out for Habbaniya (1948) and Gibraltar (1948) up to 150 mb., and for Auckland (1948), where although sufficient regularity was attained only to 250 mb. the results were so closely parallel to Gibraltar that extrapolation to 150 mb. was justifiable. For a station close to 30° latitude the best available to the writer was Norfolk Island (July 1944 to June 1945) extending as far as 350 mb. The mean annual fluxes are given in Table 3.

TABLE 3  
ANNUAL MEAN VALUES OF  $\overline{v'u'}$  AND TOTAL POLEWARD EDDY FLUX

Station	$\overline{v'u'}$ at (mb.)								Eddy Stress	Total Ang. Mom. Flux
	950	900	700	500	400	300	200	150		
Habbaniya Lat. 33° 22' N.		-0.8	-0.1	1.4	2.2	3.7	3.5	0.8	+1.0	+5.6
Gibraltar Lat. 36° 09' N.		-0.5	0.3	1.1	1.7	2.6	4.6	2.1	+1.0	+5.4
Auckland Lat. 36° 51' S.	-0.9	-1.1	0.2	0.8	1.4	2.2	[4.2]	[1.8]	+0.7	+3 to +4
Norfolk Island Lat. 29° 03' S.	-0.4	-0.5	-0.2	0.0	-0.2					
			$\times 10^5 \text{ cm.}^2/\text{sec.}^2$						$\times 10^8$ dyne/cm.	$\times 10^{23}$ g. cm. <sup>2</sup> /sec. in 1 year



The detailed figures are of considerable interest, but the main concern of the present study is with the final column and its relation to (4). Limited though the Norfolk Island values are, they show no trend suggesting that the total flux there could approach the positive magnitudes attained at the other stations. These in turn, adopting 4 to  $5 \times 10^{33}$  as a representative value, amount to only about one-third of the total annual flux required.

The standing-eddy flux of angular momentum is to be represented by

$$\frac{R}{g} \int_0^{2\pi} \int_0^{p_0} \bar{v}_\psi \bar{\Omega}_\psi \cos \varphi \, dp d\psi,$$

where the symbols with suffix  $\psi$  refer to mean seasonal values, which vary with longitude  $\psi$ . To evaluate this from actual winds would be an enormous task even if the data were available, which is not the case, and here alone in the present study the geostrophic approximation must be invoked. Using the charts of Brooks *et al.* (7), the evaluation was carried out for latitude  $30^\circ \text{N.}$ , for which the charts extended to 130 mb. Up to 300 mb., at which the charts for  $30^\circ \text{S.}$  terminated, there was no significant difference between the two hemispheres. The total annual flow poleward amounted to

$$+0.4 \times 10^{33} \text{ g. cm.}^2/\text{sec.},$$

a manifold disparity with (4) which cannot be attributed to lack of sufficient detail in the charts employed.

This evidence does not support the contention that the flow of  $\Omega$  across the high pressure belt is encompassed almost entirely by horizontal rotations. It may be remarked that similar evaluations by the present writer for latitudes  $50^\circ(15)$  and  $10^\circ$  (unpublished data) have indicated no similar discrepancy, and the view that  $\Omega$ -transfer across these latitudes is mainly effected by horizontal eddies is therefore tenable.

It is of interest here to add the results of a parallel evaluation for Norfolk Island of the annual mean poleward eddy-flux of heat in sensible and latent form. These were:

Sensible heat	..	..	..	$+0.5 \times 10^7$ cal./cm./min.
Latent heat	..	..	..	$+0.9 \times 10^7$ cal./cm./min.

of the same magnitude and sign as the total heat flow across latitude  $30^\circ$  required to maintain the annual balance between incoming and outgoing radiation. The indications that the horizontal rotations in these latitudes, despite their relative inefficiency in momentum transport, are yet significant agents in heat transport provide quantitative support for the thermodynamic arguments of the previous section.

## VI. EVIDENCE FOR THE DRIFT

From equation (6) and the known distribution of  $\bar{u}$  it is clear that any evidence for the existence of a mean drift will itself constitute evidence for the importance of toroidal circulations in transporting momentum across the high pressure belt. In a way this evidence is self-contained, for the presence of a westerly jet-stream which at times attains 150–200 m.p.h., and the absence of

reports of easterly winds approaching equivalent strength at the same (upper tropospheric) levels, do themselves suggest a resultant drift in the sense invoked.

It is possible to derive direct evidence of cross-isobaric wind components by means of slow-rising or zero-lift balloons which can be tracked for long periods. In a paper read before the Seventh Pacific Science Congress (New Zealand 1949), Dr. A. F. Spilhaus described observations of the tracks of balloons flying with ballast so controlled as to remain between 300 and 100 mb. In private correspondence Dr. Spilhaus has reported: "The results are somewhat subjective but are consistent enough to convince me that there is a very real cross-isobar component towards the poles at the altitudes at which we made constant level balloon flights . . . at levels between 100–300 mb. from higher to

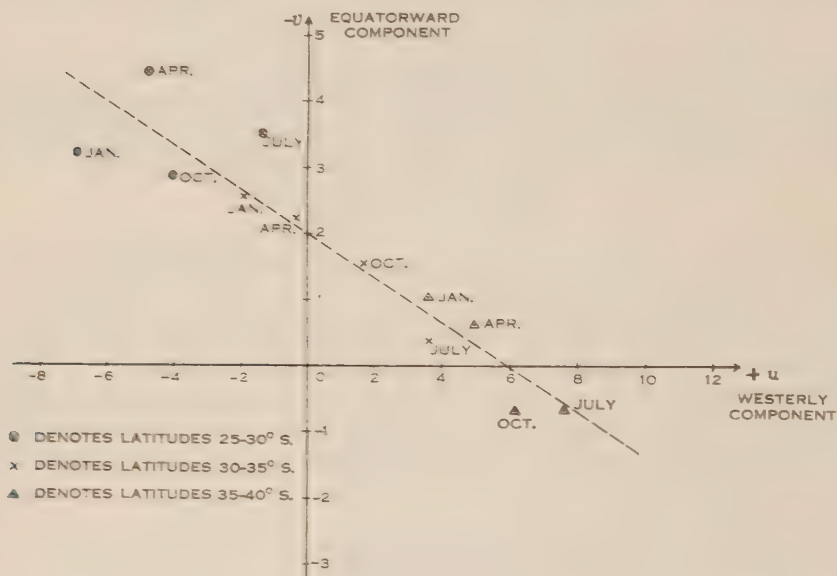


Fig. 1.—Annual mean components of surface wind.

lower pressure in a range of latitudes from 30 to about 45°." More such flights, in all situations and longitudes, are desirable before the order of magnitude of the mean drift can be directly assessed.

On the other hand the counter-drift implied by (7) could be detected and assessed if it were to occur, at least in part, in layers close to the surface. With this possibility in mind, a study was made of the surface winds over the oceans of the Southern Hemisphere from the British Admiralty Charts already referred to. The mean surface wind vector was determined for each of the zones 25–30° S., 30–35° S., 35–40° S., for the months of January, April, July, and October, and in Figure 1 is plotted the equatorward ( $-v$ ) against the westerly ( $-u$ ) component for each of the twelve vectors so derived. The calculated regression of  $v$  on  $u$ , shown by the broken line, was

$$v \text{ (m.p.h.)} = +0.33 u - 2.0.$$



The intercepts are statistically significant and the value of that on the  $v$ -axis, i.e. 2 m.p.h. towards the equator, establishes a counter-drift of the right order of magnitude. A useful by-product of this investigation is to fix at about  $32^\circ\text{S}$ . the latitude dividing mean annual easterly from westerly surface winds. The study was repeated for the oceans of the Northern Hemisphere with essentially the same results, the indications being somewhat less clear owing to the greater irregularity of pattern, made manifest in a greater scatter of individual points from the regression line.

Local mean values of  $\bar{v}$  were automatically derived in the course of the flux evaluations of Section V. The average mean values for 1948 for Habbaniya and Gibraltar (some  $50^\circ$  apart in longitude) were :

Pressure .. ..	900	700	500	400	300	200	150 mb.
Mean poleward drift ..	-100	+60	+120	+120	+140	+170	+150 cm./sec.

At first sight these figures appear to provide striking direct support for the drift, but too much weight should not be attached to local values of  $\bar{v}$ , which are largely governed by position relative to the semi-permanent pressure and thermal systems of the region. On the other hand the fact that the two stations are separated by about one-half the length of the standing long waves does add weight to this evidence.

Indirect evidence giving further support for the existence of the mean drift is afforded by the dryness of the air in the lower stratosphere of temperate latitudes(16). The low frost points there measured are comparable with those of the upper troposphere of the tropics, and there does not appear to be any other explanation as probable as that of a net poleward drift at these levels. It may be remarked that the mean isentropic surfaces are so disposed that air drifting poleward along these surfaces from the upper troposphere in low latitudes would appear in the lower stratosphere of the temperate zone.

Thus the evidence for the drift is concordant. From (6) and (8) it must be recognized that the relation between the drift and the index is an intimate one, though the precise nature of the relation may eventually prove more complex than has so far been outlined. One such possible complexity appears in the next section where the theory is submitted to its final test, a comparison of the sequence of events predicted from it with that occurring in real situations.

## VII. SYNOPTIC PHENOMENA

The theory put forward pertains essentially to a symmetrical atmosphere, but the cycle depends at one stage on the realization of an instability and it is not likely that in practice this stage will be reached simultaneously in all longitudes. Since there are time-lags in the operation of various parts of the mechanism and since, also, the underlying surface of the earth shows a considerable lack of homogeneity with obstacles in the form of mountain ranges extending to great heights, it is clear that zonal symmetry must represent an over-idealization.

In practice the ideal model might be approximated by a pattern of waves, of which nodes or troughs tend to be anchored in some longitudes and which in

other regions take on a progressive form. It was from the intensive studies of such waves, in the fields of synoptic and regional meteorology, that the existence of the jet-stream came to be recognized. The present paper assigns to the jet-stream an active role in the more general phenomenon, and it is therefore of fundamental interest to determine to what extent the main features of the symmetrical model appear in the wave oscillations. Since the formal dynamics of the symmetrical and wave problems differ only in the appearance of advective terms in the latter, the extent of agreement between the indications of the present theory and the observed wave phenomena will afford not only a test of the significance of the theory, but also a broad measure of the importance of advective effects in the wave cycle.

Evidence has been given for the drift and there can be little doubt of the reality of the other broad features, i.e. the attainment of dynamic instability, the oscillations of the index with a period of the order 10–20 days, and the fluctuations of the jet stream axis which now appear as the “meanderings of a river” from which the name originated. The theory provides further pointers of whose correctness in practice we are at present less sure, and which appear deserving of further synoptic examination. Among these may be mentioned:

(i) A positive correlation between the latitude of the jet stream axis and the value of the surface index.

(ii) A tendency for maximum speed in the jet stream at times when the axis is not far removed from the latitude of the mean high pressure belt. There are two conflicting influences at work here, since the drift itself would tend to establish the maximum speed at the poleward limit of its influence whereas, as discussed in Sections III and IV, frictional and thermodynamical influences will tend to realize the maximum on the equatorward side. A more crucial test of the applicability of the present theory to wave phenomena might be found in (iii).

(iii) The occurrence of dynamical instability predominantly between the mean high pressure belts.

(iv) A tendency for instability to be associated with the extreme (low) latitude position of the jet stream axis.

(v) Onset or increase in the drift as a herald of increasing index and poleward movement of the jet stream axis.

(vi) As between one layer and another, the upper-levels in low latitudes and the lower levels in middle latitudes appear to play the more active role in the cyclic process.

(vii) As between one season and another, the jet stream phenomena will be most marked at seasons when the index fluctuations are strongest, and vice versa.

The only statistical investigation of short-term changes capable of throwing light on (i) appears to be that of Willett(17), and in this work some significant synoptic correlations may be somewhat obscured on account of the use of five-day mean values. Willett found, among other interesting associations, that a low index is characterized by a relatively strong poleward temperature gradient



and by relatively strong westerlies aloft with a tendency to be displaced equatorward of their mean position. Of special interest to the present theory was a consistently negative correlation between the speed of the strongest westerly winds at 10,000 feet and their latitude of occurrence, a correlation which in view of the thermal contrast between latitudes may be expected to increase significantly with height.

A summary of the results of the first extensive jet stream investigation has been published by meteorologists of Chicago University(18), and provides general though perhaps not conclusive evidence that (ii), (iii), and (iv) are features of the long wave cycles.

No evidence on (v) can be expected at present, but the suggestion that the onset of instability and drift should precede a radical change in local weather type is one which is thought to merit further investigation: so also is (vi) although it is not at present clear how the relatively active and passive roles might be made manifest.

There is some indication(8) that index fluctuations are more marked and regular in winter than in summer. If this is established it would confirm (vii) for there can be no doubt of the intensification of the jet stream during the winter season. So marked is this intensification that, although little discrimination between seasons has been made in the present text, the theory must be regarded as having main application to the winter hemisphere.

### VIII. THE MIXING LENGTH APPROACH

An earlier theory of the planetary circulation(11) is based on lateral interchange mechanisms by means of the common mixing length concepts. The theory of drift does not necessarily rest on such concepts. The two theories are, however, not essentially at variance for, if certain limitations of the mixing length formalism are overlooked, the drift appears as a natural concomitant of mixing and the present theory as an extension of the earlier one.

A parcel originating in latitude  $\varphi - \Delta\varphi$  which conserves the average angular momentum of that latitude as it crosses latitude  $\varphi$  with poleward velocity  $v'$  would effect a flux of angular momentum of amount

$$\rho v' \bar{\Omega}_{\varphi - \Delta\varphi} = \rho v' \bar{\Omega}_{\varphi} - \rho v' \Delta\varphi \cdot \frac{\partial \bar{\Omega}}{\partial \varphi}$$

per unit area. Using bars to denote means over a long period and over the whole latitude zone, the mean flux across latitude  $\varphi$  per unit depth of layer is then

$$2\pi\rho \left( \overline{K\bar{v} \cos \varphi} \cdot \bar{\Omega}_{\varphi} - K \cos \varphi \cdot \frac{\partial \bar{\Omega}}{\partial \varphi} \right),$$

where  $K$  is written for  $\overline{Kv' \Delta\varphi}$ , the usual turbulent interchange coefficient. This expression differs from that commonly encountered in so far as it recognizes that a systematic through-drift,  $\bar{v}$ , would provide an additional contribution to the flux.

Setting  $K$  independent of latitude, the condition for no accretion of angular momentum in any zone allows the sign and magnitude of  $v$  under these assumptions to be assessed from

$$\bar{v} = \frac{K \frac{\partial}{\partial \varphi} \left( \cos \varphi \frac{\partial \bar{\Omega}}{\partial \varphi} \right)}{R \frac{\partial}{\partial \varphi} (\bar{\Omega} \cos \varphi)} \dots \dots \dots (11)^*$$

Taking  $\zeta$ , the absolute vorticity, rather than  $\Omega$  as the conserved entity would lead, by a similar analysis, to

$$\bar{v} = \frac{K \frac{\partial}{\partial \varphi} \left( \cos \varphi \frac{\partial \bar{\zeta}}{\partial \varphi} \right)}{R \frac{\partial}{\partial \varphi} (\bar{\zeta} \cos \varphi)} \dots \dots \dots (12)$$

in place of (11). In practice the figures for  $v$  determined from (12) would be less secure than those from (11) since  $\zeta$  represents a  $\bar{u}$ -derivative of one order higher than  $\bar{\Omega}$  so that inaccuracies in  $\bar{u}$  would lead to greater uncertainties in the result.

Using wind data(6) for the most favourable season, winter, and level, 12 km., a skeleton evaluation of (11) and (12) is given in Tables 4 and 5 respectively. Defant's(19) figure of  $10^{11}$  cm.<sup>2</sup>/sec. for  $K$  is used as representing the largest value it is at all reasonable to assume. By comparison with (8), a mean figure for a year and for a deep layer, it is seen that the resulting figures for  $v$  are of the same sign and order of magnitude as that assigned to the drift.

While this agreement and possible extension of the mixing length theory are worth noting, it is preferred that the main weight for the drift-index theory should rest on the considerations of the earlier sections since the application of the mixing-length formalism to this scale of phenomenon is insecure. The result that the flux is proportional to, or even of the same sign as, the gradient of the entity considered has been discountenanced by direct determinations(15). Nor does the theory apply unless either the distribution of  $u$  with  $\varphi$  can be regarded as sensibly steady, which is only the case for a few days at a time(8), or unless the scale of mixing were independent of the current values of  $u$ ,  $\partial u / \partial \varphi$ , etc. which is excluded by the study(10) of the stable and unstable perturbation orbits. A mixing-length theory employing the mean annual or mean seasonal  $\bar{u}$ ,  $\varphi$  distribution appears to suffer from the same inadequacies as would a theory of turbulent heat conduction from the surface based on mean daily vertical temperature distributions.

\* From (5) and (11), with  $\bar{u}=0$ , it follows that

$$v = \frac{2K}{3R} (\cot \varphi - 2 \tan \varphi),$$

which is positive from the equator to  $\varphi \approx 35^\circ$  and negative beyond. Despite anomalies at the equator and poles this solution is of physical interest to theories of the origin of planetary circulations, in indicating that drift and confluence might be initiated by a purely dynamic stimulus.

TABLE 4  
COMPUTATION OF  $\bar{v}$  FROM EQUATION (11) USING WILLET'S WINTER DATA AT 12 KM.

$\varphi$	$\bar{u}$ (cm./sec.)	$\bar{\Omega}$ (cm. <sup>2</sup> /sec.)	$\frac{\partial \bar{\Omega}}{\partial \varphi}$ (cm. <sup>2</sup> /sec.)	$R\bar{\Omega} \cos \varphi$ (cm. <sup>3</sup> /sec.)	$K \frac{\partial}{\partial \varphi} (\cos \varphi \frac{\partial \bar{\Omega}}{\partial \varphi})$ (cm. <sup>4</sup> /sec. <sup>2</sup> )	$R \frac{\partial}{\partial \varphi} (\bar{\Omega} \cos \varphi)$ (cm. <sup>3</sup> /sec.)	$\bar{v}$ (cm./sec.)
15°	+1720	28.66 × 10 <sup>12</sup>	-4.7 × 10 <sup>12</sup>	17.63 × 10 <sup>21</sup>		-8.1 × 10 <sup>21</sup>	
17½°	+3500	28.24 × 10 <sup>12</sup>	-12.4 × 10 <sup>12</sup>	16.91 × 10 <sup>21</sup>	-8.0 × 10 <sup>24</sup>	(-11.2 × 10 <sup>21</sup> )	+700
20°	+4990	27.16 × 10 <sup>12</sup>	-21.3 × 10 <sup>12</sup>	15.67 × 10 <sup>21</sup>	-8.6 × 10 <sup>24</sup>	(-14.2 × 10 <sup>21</sup> )	+500
22½°	+5650	25.30 × 10 <sup>12</sup>	-33.1 × 10 <sup>12</sup>	13.96 × 10 <sup>21</sup>	-10.4 × 10 <sup>24</sup>	(-16.9 × 10 <sup>21</sup> )	+450
25°	+4930	22.41 × 10 <sup>12</sup>	-41.7 × 10 <sup>12</sup>	11.69 × 10 <sup>21</sup>	-6.0 × 10 <sup>24</sup>	(-19.6 × 10 <sup>21</sup> )	
27½°						(-22.8 × 10 <sup>21</sup> )	
30°						(-26.0 × 10 <sup>21</sup> )	
32½°						(-27.5 × 10 <sup>21</sup> )	
35°						(-29.0 × 10 <sup>21</sup> )	
37½°							
40°							

( ) denotes interpolated values.

TABLE 5  
COMPUTATION OF  $\bar{v}$  FROM EQUATION (12) USING WILLET'S WINTER DATA AT 12 KM.

$\varphi$	$\bar{u}$ (cm./sec.)	$\frac{\partial \bar{u}}{\partial \varphi}$ (cm./sec.)	$\bar{\zeta}$ (sec. <sup>-1</sup> )	$\frac{\partial \bar{\zeta}}{\partial \varphi}$ (sec. <sup>-1</sup> )	$R\bar{\zeta} \cos \varphi$ (cm./sec.)	$K \frac{\partial}{\partial \varphi} (\cos \varphi \frac{\partial \bar{\zeta}}{\partial \varphi})$ (cm. <sup>2</sup> /sec. <sup>2</sup> )	$R \frac{\partial}{\partial \varphi} (\bar{\zeta} \cos \varphi)$ (cm./sec.)	$\bar{v}$ (cm./sec.)
15°	+1720	+20.4 × 10 <sup>3</sup>	+1.32 × 10 <sup>-5</sup>	+2.14 × 10 <sup>-4</sup>	+8.01 × 10 <sup>3</sup>		+12.3 × 10 <sup>4</sup>	
17½°	+3500	+17.1 × 10 <sup>3</sup>	+3.18 × 10 <sup>-5</sup>	+3.20 × 10 <sup>-4</sup>	+18.7 × 10 <sup>3</sup>	+10.4 × 10 <sup>7</sup>	(+14.8 × 10 <sup>4</sup> )	+700
20°	+4990	+7.6 × 10 <sup>3</sup>	+5.98 × 10 <sup>-5</sup>	+4.21 × 10 <sup>-4</sup>	+33.8 × 10 <sup>3</sup>	+8.6 × 10 <sup>7</sup>	(+17.3 × 10 <sup>4</sup> )	+450
22½°	+5650	-8.25 × 10 <sup>3</sup>	+9.66 × 10 <sup>-5</sup>	+3.82 × 10 <sup>-4</sup>	+51.9 × 10 <sup>3</sup>	-6.0 × 10 <sup>7</sup>	(+19.0 × 10 <sup>4</sup> )	
25°	+4930	-23.3 × 10 <sup>3</sup>	+13.00 × 10 <sup>-5</sup>		+65.6 × 10 <sup>3</sup>		(+20.7 × 10 <sup>4</sup> )	
27½°							(+18.2 × 10 <sup>4</sup> )	
30°							(+15.8 × 10 <sup>4</sup> )	
32½°								
35°								
37½°								
40°								

( ) denotes interpolated values.



## IX. ACKNOWLEDGMENTS

Thanks are due to the directors and staffs of the Meteorological Offices of Great Britain and New Zealand for the provision of upper air data, and to Miss R. I. Thomas for the heavy work of extraction and computation required at various stages of the investigation.

## X. REFERENCES

- (1) BRUNT, D.—“Physical and Dynamical Meteorology.” 2nd Ed. (Cambridge Univ. Press, 1939.)
- (2) JEFFREYS, H.—*Quart. J.R. Met. Soc.* **52**: 85 (1926).
- (3) ROSSBY, C. G., and STARR, V. P.—*J. Met.* **6**: 288 (1949).
- (4) BRUCH, H.—*Veröff. Inst. Meeresk. Univ. Berl.* Neue Folge A **38**: 5 (1940).
- (5) ROSSBY, C. G., and MONTGOMERY, R. B.—*Pap. Phys. Oceanogr. Met., Camb. Mass.* **4** (3) (1936).
- (6) WILLETT, H. C.—“Descriptive Meteorology.” (Academic Press, 1944.)
- (7) BROOKS, C. E. P., *et al.*—Air Minist. (Lond.) Geophys. Mem. No. 85 (in press).
- (8) NAMIAS, J.—“Extended Forecasting by Mean Circulation Methods.” (U.S. Weather Bureau, 1947.)
- (9) EADY, E. T.—*Tellus* **1**: 33 (1949).
- (10) VAN MIEGHEM, J.—*Inst. R. Met. Belg. Mem.* **23** (1946).
- (11) ROSSBY, C. G.—*Bull. Amer. Met. Soc.* **28**: 53 (1947).
- (12) ROSSBY, C. G.—“Climate and Man.” p. 599. (U.S. Dep. Agric., 1941.)
- (13) JEFFREYS, H.—*P.V. Mét. Un. Géod. Géophys. Int.*, Lisbon **1933**: 219 (1933).
- (14) HESS, S. L.—*J. Met.* **5**: 293 (1948).
- (15) PRIESTLEY, C. H. B.—*Quart. J.R. Met. Soc.* **75**: 28 (1949).
- (16) DOBSON, G. M. B., BREWER, A. W., and CWILONG, B. M.—*Proc. Roy. Soc. A* **185**: 144 (1946).
- (17) WILLETT, H. C.—*Trans. Amer. Geophys. Un.* **29**: 803 (1948).
- (18) CHICAGO UNIVERSITY, DEPARTMENT OF METEOROLOGY.—*Bull. Amer. Met. Soc.* **28**: 255 (1947).
- (19) DEFANT, A.—*Geogr. Ann., Stockh.* **3**: 209 (1921).

# GALACTIC RADIATION AT RADIO FREQUENCIES

## I. 100 Mc/s. SURVEY

By J. G. BOLTON\* and K. C. WESTFOLD\*

[*Manuscript received October 21, 1949*]

### *Summary*

This paper contains an account of a survey of galactic radiation at a frequency of 100 Mc/s. An aerial array with a  $17^\circ$  beamwidth, on an equatorial mounting, was used to plot the distribution of intensity over the section of the celestial sphere between Declination  $+30^\circ$  and  $-90^\circ$ . The method of eliminating the effect of the aerial polar diagram from the observations is described and the final distribution, expressed in terms of equivalent black-body temperature, is presented in galactic coordinates on a series of equal-area charts.

## I. INTRODUCTION

### (a) *Historical*

Since the discovery by Jansky(1) in 1932 of galactic radiation at radio frequencies, surveys of intensity distribution over the celestial sphere have been made by a number of workers. The most accurate are those due to Reber(2, 3) at 160 and 480 Mc/s. and Hey, Parsons, and Phillips(4) at 64 Mc/s. The present survey, covering the region between galactic longitudes  $140^\circ$  and  $40^\circ$  is of particular importance as it includes the portion between the galactic centre ( $l=330^\circ$ ) and Monoceros ( $l=200^\circ$ ), which cannot be observed from northern latitudes. The frequency of 100 Mc/s. used in this survey is sufficiently close to 64 Mc/s. used by Hey *et al.* and 160 Mc/s. used by Reber to enable a complete picture of the noise distribution over the whole celestial sphere to be constructed. Such a picture will undoubtedly be of value in studying correlation between radio-frequency and optical data and possibly in deducing the form of the Galaxy.

### (b) *Temperature Unit*

The intensity contours of previous surveys have been plotted in units of received power, e.g. watts cm.<sup>-2</sup> (c/s.)<sup>-1</sup> steradian<sup>-1</sup>. Theoretical writers, however, seem to prefer results expressed in terms of the equivalent black-body temperature, owing to the possible relationship with the electron temperature of interstellar matter. The present writers have adopted the temperature unit and also the concept of aerial temperature for making observations. The calibration of the receiver may conveniently be made in terms of the thermal agitation noise generated in an artificial load.

\* Division of Radiophysics, C.S.I.R.O., University Grounds, Sydney.

Owing to the relatively large beamwidth of the aerial system the observed aerial temperatures represent weighted averages of the equivalent temperature over considerable areas of the celestial sphere. In Section V details are given of the method of analysis used to determine the distribution of equivalent temperature from the observed distribution of aerial temperature and a knowledge of the aerial polar diagram. An approximate mathematical reduction was used first, followed by a trial-and-error method somewhat similar to that adopted by Hey, Parsons, and Phillips.

## II. EQUIPMENT

### (a) Aerial System

The aerial system used in this investigation is shown in Plate 1. The array of nine Yagis in three groups of three, one wavelength apart, is fixed on

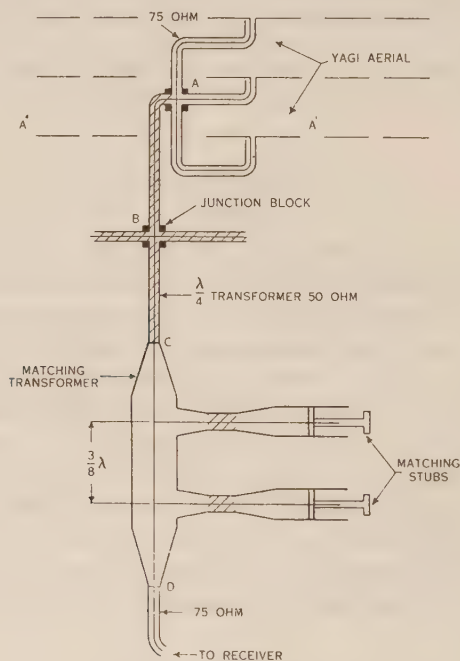


Fig. 1.—Details of aerial matching system.

an equatorial mounting. A “split” in the counterweight arm between the declination and polar axes enables part of this arm to be rotated with respect to the rest. With the counterweight arm horizontal and the split correctly adjusted, the declination axis can be used as an azimuth axis. (The aerial is shown in such a position in the inset to Plate 1.) Selsyn repeaters and remotely-controlled electric motors are fitted to all three axes.

### (b) Aerial Matching

The matching arrangements of the aerial array are shown in diagrammatic form in Figure 1. The Yagis are individually matched to an impedance of



75 ohms. Each set of three is connected together at junction blocks  $A$ ,  $A'$ ,  $A''$ . The impedance of these sets at the junction blocks is 25 ohms. The three blocks are joined to a further junction block,  $B$ , by quarter-wavelength cable transformers of 50-ohm impedance. The impedance of a single set of aerials seen from  $B$  is thus 100 ohms and the three sets in parallel,  $33\frac{1}{3}$  ohms. Junction block  $B$  is joined to an air-cored double-stub transformer by a quarter-wavelength 50-ohm cable transformer  $BC$ . At  $C$ , therefore, the impedance of the whole array is 75 ohms. The double-stub matching transformer, which is housed in an enclosure on the counterweight arm of the aerial mounting, is used to remove any reactive components introduced at junction points etc. This transformer is connected to the receiver by a 75-ohm cable. The enclosure is maintained at a temperature of about 70 °F. by electric elements to prevent condensation on the transformer and matching stubs.

For making observations of galactic noise, the output of the receiver detector was amplified by a D.C. amplifier. The output of this amplifier was fed on to an Esterline-Angus recording milliammeter. The zero on this scale was set by substituting a matched load for the aerial and the scale calibrated with a diode noise generator, directly in units of temperature. Although each component of the system was carefully matched to 75 ohms, a further check was carried out before each day's observations. Additional short cable lengths were interposed between receiver and aerial, or receiver and noise generator. The absence of a change in reading on the milliammeter was taken as an indication of satisfactory matching.

### (c) *Aerial Polar Diagram*

Measurements of the aerial polar diagram were made using a dipole as a signal source. This dipole was placed close to the ground in order to minimize the effect of ground reflections. Various sites were used at distances between 100 and 300 yards from the aerial to eliminate the possibility of stray reflections from surrounding objects. The polar diagram (in terms of relative power) was plotted automatically on a recording milliammeter as the aerial was turned past the signal source. Measurements were made with the individual Yagis of the system interchanged to check for differences in the electrical lengths of their feeders or other defects.

The aerial diagram was found to consist of a main lobe,  $17^\circ$  between the half power points, and seven subsidiary lobes. Of the total power 67 per cent. was concentrated in the main lobe and about 5 per cent. in each of the first five of the subsidiary lobes listed below.

(1) A back lobe, the same shape as the main lobe.

(2, 3) Two side lobes centred in the plane of the array and in the direction of polarization, approximately the same shape as the main lobe.

(4, 5) Two side lobes centred in the plane of the array and at right angles to the direction of polarization, approximately the same shape as the main lobe.

(6, 7) Two comparatively weak lobes centred  $120^\circ$  from the main lobe in the plane of polarization.

The main lobe was found to be almost a circular cone as is shown in Figure 2; for the purpose of analysis it was assumed to be circular. The gain of the array with respect to an isotropic radiator was found by integration to be 71.

### III. THEORETICAL

#### (a) Aerial Relationships

The intensity  $I(f; l, b)$  of linearly polarized radiation in the range of frequency  $(f, f+df)$  from a direction specified by coordinates of galactic longitude  $l$  and latitude  $b$  may be conveniently expressed in terms of the temperature,  $T_e$ , of black-body radiation having the same intensity.

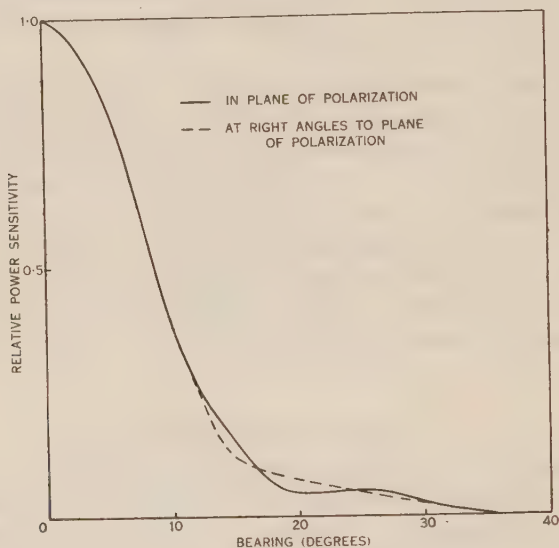


Fig. 2.—Polar diagram of the main lobe of the aerial in and perpendicular to the plane of polarization.

The relation

$$I(f; l, b) = \frac{k}{c^2} f^2 T_e(f; l, b) \dots\dots\dots (1)$$

thus defines the equivalent temperature. The contribution to the power delivered to the aerial output terminals in the frequency band  $df$  of the radiation received from the elementary solid angle  $d\omega$  about the direction  $(l, b)$  is

$$F(\theta\varphi) A I(f; l, b) df d\omega,$$

where  $F(\theta\varphi)$  represents the power sensitivity of the aerial and  $A$  its effective area for normal incidence. The polar angle  $\theta$  of the direction  $(l, b)$  is measured from an axis through the centre of the aerial perpendicular to the plane of the array in the direction  $(l_0, b_0)$  and the azimuth angle  $\varphi$  is measured from the direction of polarization in the plane of the aerial (see Fig. 3).

The aerial output power from the whole celestial sphere is

$$A df \int F(\theta, \varphi) I(f; l, b) d\omega \dots\dots\dots (2)$$

Burgess(5) has shown that the power delivered from an aerial in equilibrium with radiation of temperature  $T$  to a matched load is the same as that from a resistor at the same temperature, viz.  $kT\Delta f$ .

Thus if  $T_a$  is the aerial temperature, (1) and (2) give

$$\begin{aligned} T_a(f; l_0, b_0) &= \int F(\theta\varphi) \frac{Af^2}{c^2} T_e(f; l, b) d\omega, \\ &= \int K(\theta\varphi) T_e(f; l, b) d\omega \dots\dots\dots (3) \end{aligned}$$

The term  $K(\theta\varphi)$  may be found from the aerial polar diagram; the normalized value  $\int K(\theta\varphi) d\omega = 1$ .

### (b) Receiver Relationships

The aerial temperature  $T_a$  is measured in terms of the power delivered to a matched receiver. The effect of power losses in the line between the aerial output and receiver input terminals is represented by the factor  $\alpha$ . Thus, of the power delivered from the aerial,  $\alpha kT_a \Delta f$  is delivered to the receiver input

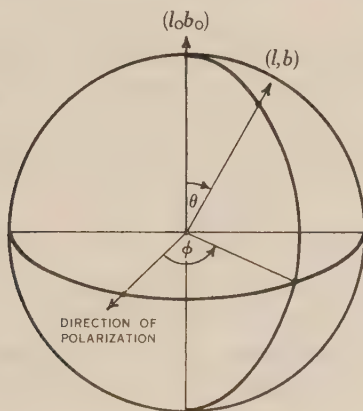


Fig. 3.—The polar coordinates  $\theta, \varphi$  associated with the aerial.

terminals and  $(1-\alpha)kT_a \Delta f$  is absorbed by the line. The line, however, is in equilibrium at the ambient temperature  $T_0$  and therefore at the same time contributes power  $(1-\alpha)kT_0 \Delta f$  to the receiver input. Hence the total power available at the receiver input terminals is

$$k\{T_0 + \alpha(T_a - T_0)\} \Delta f.$$

If  $N$  is the noise factor\* of the receiver, whose bandwidth is  $\Delta f$ , its own contribution to the power output is  $(N-1)kT_0 \Delta f$ . The total power is thus

$$P = N k T_0 \Delta f + \alpha k (T_a - T_0) \Delta f.$$

If an artificial load at the ambient temperature be substituted for the aerial the total power output is

$$P_0 = N k T_0 \Delta f.$$

\*  $N$  may be found by using an artificial source of noise such as a saturated diode or thermistor.



Then

$$\frac{T_a - T_0}{T_0} = \frac{N}{\alpha} \frac{P - P_0}{P_0}.$$

#### IV. METHOD OF OBSERVATIONS

##### (a) *Equatorial Axis*

In making observations the aerial was turned from horizon to horizon along arcs of constant declination. Sweeps were made from east to west and west to east, the declination being changed by  $3^\circ$  after every east-west-east sweep. Each sweep occupied about five minutes and a recorder speed of three inches per minute was used, giving a resolution of  $12^\circ$  per inch of recorder chart. The Right Ascension was obtained by comparing a sidereal clock and the hour-angle repeater dial of the aerial. The observer pressed a key at integral hours of Right Ascension which operated a second pen on the recorder.

Observations were made at different times of the day and night to eliminate horizon effects and solar radiation. Generally the same aerial temperature was

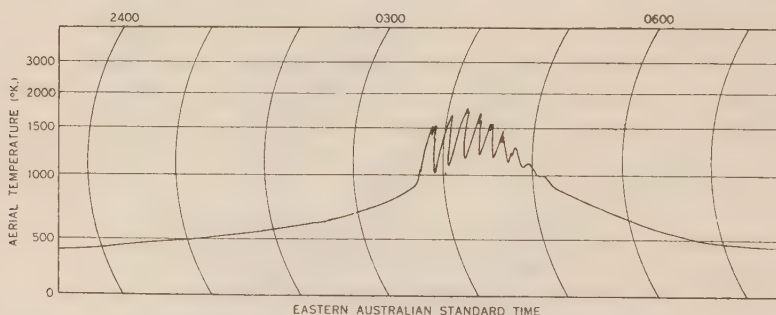


Fig. 4.—Interference pattern of the discrete source in Cygnus superimposed on the local diffuse maximum in galactic noise.

observed when looking at the same point on the celestial sphere. In certain cases, however, dual sets of values were obtained, depending on whether the main body of the Galaxy was above the horizon and contributing through the subsidiary lobes of the aerial.

##### (b) *Azimuth Axis*

A further survey was made with the aerial on its azimuth mounting (Section II (a) and Plate 1). The aerial was left for 24 hours on a fixed easterly bearing overlooking the sea. From this survey it was possible, owing to interference from sea reflection, to distinguish localized source regions of angular diameter less than  $1^\circ$ . This is well illustrated in Figure 4 which depicts a strong discrete source in Cygnus(6) superimposed on a local diffuse maximum of galactic noise. The effect of this source on the polar-axis survey is to increase the aerial temperature by about  $150^\circ\text{K.}$  in the direction of the source. The less powerful sources in Virgo, Taurus, and Centaurus(7), raise the aerial temperature by about  $15^\circ\text{K.}$  The effect of these and a number of other smaller sources was regarded as negligible.

# V. RESULTS AND ANALYSIS

## (a) *Reduction of Observations*

The records obtained from the observations described above were marked at 100° temperature intervals and the corresponding points plotted in equatorial coordinates on a rectangular chart. Smooth contours were then drawn through this network of points. In cases of dual readings, the values corresponding to observations near the meridian were used in order to simplify correction for the effects of subsidiary lobes of the aerial diagram.

These contours were transferred to a series of equal-area charts in galactic coordinates shown in Figures 5A and 5B. Galactic coordinates were selected owing to the concentration of high temperature contours about the galactic equator, and equal-area charts were chosen to coincide approximately with sections of the celestial sphere. Contours representing the aerial polar diagram then had the same shape on the charts wherever the aerial was centred. These contours were approximately circular on the celestial sphere so that difficulties of orientation in transferring from the equatorial coordinates of observation to galactic coordinates did not arise.

The contours shown in Figure 5 near galactic longitude 40° have been corrected from the original observations for the effect of the strong discrete source in the constellation of Cygnus. The intensity of this source was found from records of the secondary survey using the sea reflection technique.

## (b) *Analysis*

As pointed out previously the observed aerial temperatures represent averages of the actual equivalent temperatures of the celestial sphere over the aerial diagram. The problem of finding the true distribution of equivalent temperature over the celestial sphere requires the solution of equation (3) which, dropping the frequency suffix, may be written

$$T_a(l_0, b_0) = \int K(\theta\phi) T_e(l, b) d\omega, \dots\dots\dots (4)$$

where  $K(\theta\phi)$  includes both the main and subsidiary lobes of the aerial diagram.

The solution was obtained in two stages, by neglecting the subsidiary lobes at first and introducing corrections at a later stage. In addition the integral  $\int K(\theta\phi) d\omega$  was replaced by a summation represented by weighted areas on a transparent grid.

Equation (4) may be written using matrix notation

$$T_a = K T_e^* \dots\dots\dots (5)$$

Approximately

$$\begin{aligned} T_e &= K^{-1} T_a, \\ &= \{I - (I - K)\}^{-1} T_a, \\ &= \{I + (I - K) + (I - K)^2 \dots\} T_a, \end{aligned}$$

\*  $\int K(\theta\phi) d\omega$  is used to denote the integral function representing the aerial diagram,  $K$  the normalized value, and  $K$  the summation process replacing the integral.  $T_a$  and  $T_e$  refer to point temperatures and  $T_a$  and  $T_e$  to the temperature distribution.

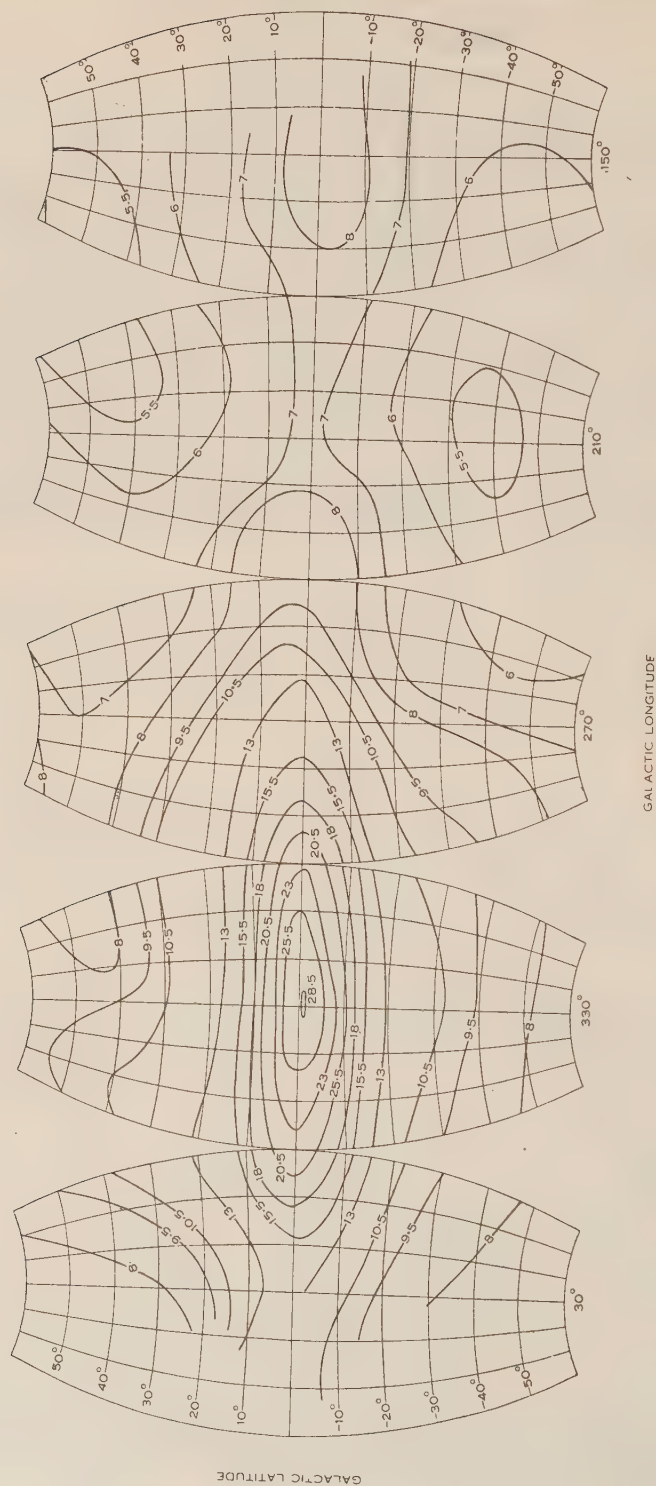


Fig. 5A.—Equal-area charts of observed aerial temperature over the celestial sphere (units of 100 °K. in galactic coordinates).



giving a sequence of successive approximations

$$\begin{aligned} T_e^{(0)} &= T_a, \\ T_e^{(1)} &= 2T_a - \mathbf{K}T_a, \\ T_e^{(2)} &= T_a + T_e^{(1)} - \mathbf{K}T_e^{(1)} \dots\dots\dots (6) \end{aligned}$$

The process represented by  $\mathbf{K}T_a$  is virtually the same as that made by the aerial system on the *actual* temperature distribution but in this approximation it is made on the *observed* distribution. The grid used in this reduction is shown in Figure 6A, and a comparison between the weighted areas of the grid and the actual aerial diagram in Figure 7. The grid was placed over a series of points forming a lattice at intervals of  $5^\circ$  in galactic longitude and latitude on the charts of aerial temperature. The mean value of the aerial temperature within

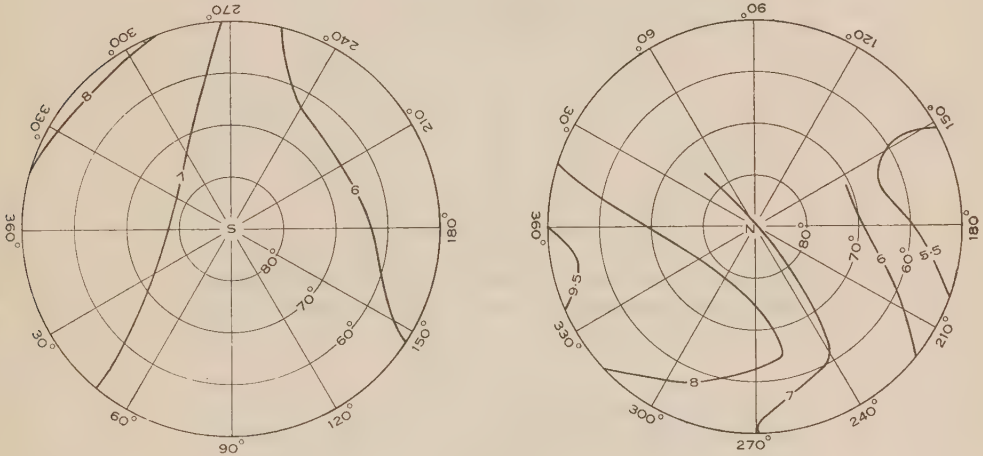


Fig. 5B

each section of the grid was multiplied by the appropriate weighting factor and divided by the sum of the weighting factors (in this case, 8) to give the value  $\mathbf{K}T_a$ . Two stages of the approximation in (6) were carried out and it was noticed that in parts of the celestial sphere where the contour gradients were not steep the values obtained for  $T_e^{(2)}$  and  $T_e^{(1)}$  differed very little from  $T_a$ . In such cases, principally near the galactic poles, the effect of the finite aerial beam was therefore negligible, and the observed aerial temperatures were equal to the equivalent temperatures. Working from these areas and using the  $T_e^{(2)}$  distribution a trial-and-error procedure was adopted. Placing the transparent grid at a point on a  $T_e^{(2)}$  contour the value of  $\mathbf{K}T_e^{(2)}$  was found and compared with the value of  $T_a$  at that point. Depending on whether the value of  $\mathbf{K}T_e^{(2)}$  was greater or less than  $T_a$ , the  $T_e^{(2)}$  contour was moved away from, or towards, the polar region. Surrounding contours were adjusted accordingly. This process was continued until the operation  $\mathbf{K}T_e^{(2)}$  at any point on the chart gave approximately the observed value of the aerial temperature. It was repeated with particular care in the regions of high contour

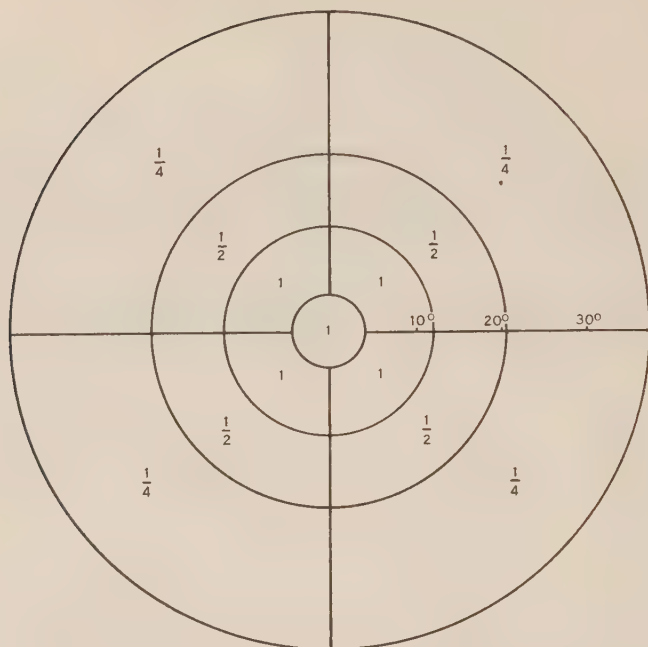


Fig. 6A.—Subdivision of main lobe into weighted areas used in first approximation (units of  $1/8$ , see Section V (b)).

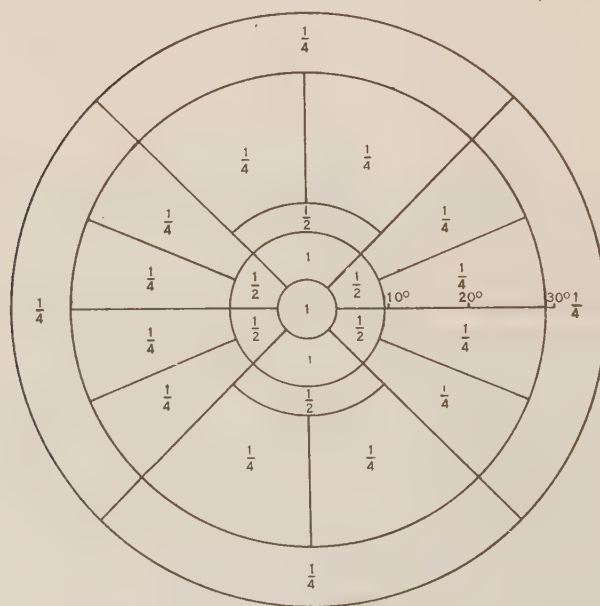


Fig. 6B.—Finer subdivision used in subsequent stages, (units of  $1/10$ , see Section V (b)).

gradient, with the grid shown in Figure 6B. This grid has finer subdivisions and the sum of the weighting factors is 10. The asymmetric arrangement of weighted areas proved convenient near the galactic equator as the temperature contours approximate to a "line source" rather than a "point source". The two arrangements are, however, equivalent. In certain regions weighting was carried out inside the areas on the grid. The final distribution (hereafter known as  $T_e'$ ) reproduced the observed distribution in most areas exactly, the maximum discrepancy between  $KT_e'$  and  $T_a$  being less than 2 per cent.

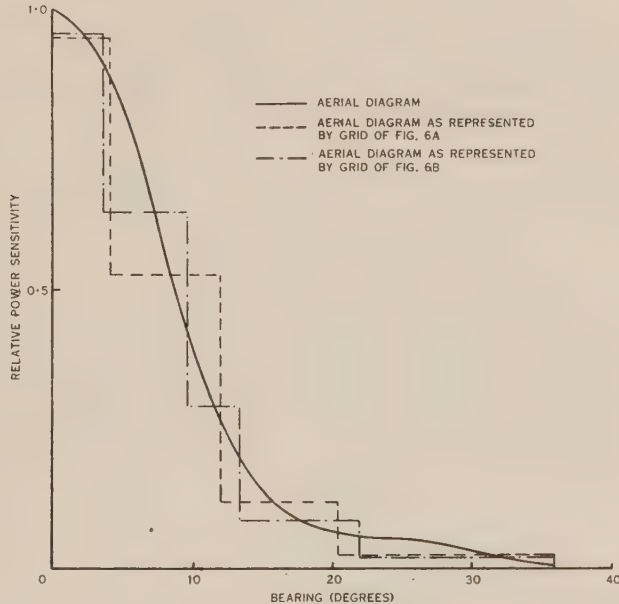


Fig. 7.—Comparison between the aerial polar diagram and the "grids" used in analysis.

### (c) Correction for Side Lobes

In the previous subsection only the effect of the main lobe of the aerial diagram has been considered. Equation (4) may be split into two parts,

$$T_a(l_0, b_0) = \int K_m(\theta\varphi) T_e(l, b) d\omega + \sum_{s=1}^7 \int K_s(\theta\varphi) T_e(l, b) d\omega, \dots\dots\dots (7)$$

where the subscript  $m$  refers to the main lobe and  $s$  to the subsidiary lobes. An approximate reduction was obtained by considering all the side lobes as seeing an average temperature  $T_{av}$ \*. The normalized values  $\int K_m(\theta\varphi) d\omega$  and

$\sum_{s=1}^7 \int K_s(\theta\varphi) d\omega$  were found from integration of the aerial diagram to be 0.67

\* Such a temperature may be found from the  $T_a$  distribution as it is independent of the aerial diagram

$$T_{av} = \frac{1}{\Omega} \int_{\Omega} T_a d\omega.$$





and  $0.33$  and  $T_{av}$  to be  $725^\circ\text{K}$ . The correction for side lobes applied to the previous distribution was

$$T_e = 1.46T_e' - 335. * \dots\dots\dots (8)$$

For observations near the meridian, four of the side lobes—1, 6, 7, and one of 4 and 5 (Section II (c))—were directed towards the ground and were considered as seeing the average temperature.† Either 4 or 5 was centred in the meridian plane at  $90^\circ$  from the main lobe and 2 and 3 were directed due east and due west on the horizon. Only when one or two of the side lobes were directed towards the galactic equator were further corrections necessary. The correction to the right-hand side of (8) can be shown to be

$$\int K_s(\theta\varphi)(T_e - T_{av})d\omega,$$

which is approximately  $K_s(T_a - T_{av})$ . The value of  $T_a$  at the required point was easily read off the chart in equatorial side lobes.

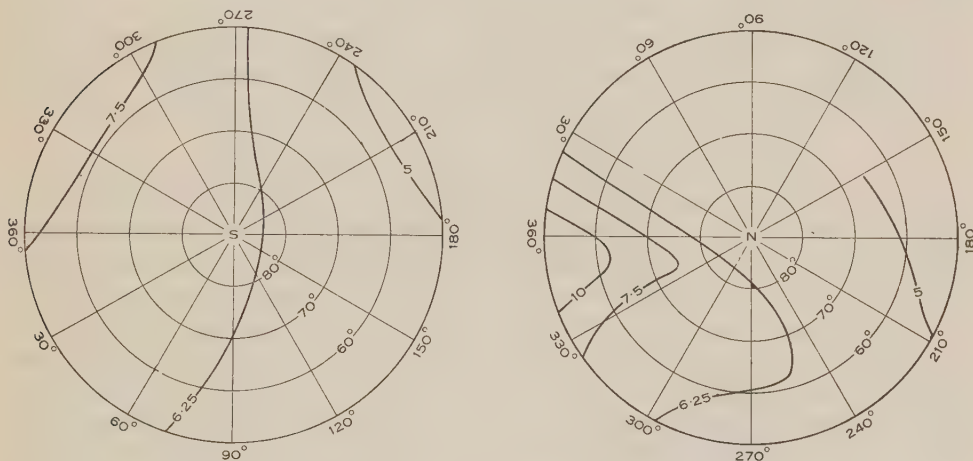


Fig. 8B

For conditions in which one or two subsidiary lobes were directed towards regions of high temperature

$$T_e = 1.46T_e' - 335 - 0.077(T_a - T_{av})_{s1} - 0.077(T_a - T_{av})_{s2} \dots\dots (9)$$

Tables were constructed for all values of  $T_e'$  and  $T_a$  which enabled rapid reduction of the  $T_e'$  contours. In cases referred to in Section V (a), where dual readings were obtained from the survey, application of the appropriate correction formulae

\* In Section V (b) we obtained the solution

$$T_a = K T_e',$$

$$T_a = K_m T_e' + \Sigma K_s T_{av},$$

$$K_m T_e' = K T_e' - \Sigma K_s T_{av}.$$

and (7) reduces to  
giving

$$T_e = \frac{T_e' - K_s T_{av}}{K_m}.$$

Approximately

† The ground (mainly wet sand) was considered as a perfect reflector. For 25 per cent. of the radiation absorbed in the ground the correction in equation (8) would be 310 instead of 335.

was found to give identical values for  $T_e$ . It can be seen that, in general, the use of equation (8) only alters the scale of the contours  $T_e'$  to  $T_e$ , whereas equation (9) alters both the scale and the shape.

The final distribution of equivalent temperature at 100 Mc/s. is shown in Figures 8A and 8B.

## VI. ERRORS

The sources of possible errors in this investigation fall into two classes, errors in calibration and errors in observation and analysis.

### (a) *Errors in Calibration*

(i) In estimating the power loss factor between the aerial output and receiver input, the matching network was considered to have a loss equal to that of an equivalent length of 75-ohm feeder. This was justified by the low standing-wave ratio ( $<1.01$ ) in the system.

(ii) The aerial was considered to have a linear response over the bandwidth of the receiver, as had the noise generator used in calibration. The maximum error in the observed aerial temperatures is estimated as less than 1 per cent.

### (b) *Errors in Observations and Analysis*

The error in a single observation of the aerial temperature at a particular point was estimated as  $\pm 50^\circ$ . This includes both reading accuracy and possible shifts in the zero of the temperature scale calibration. The final aerial temperature contours were drawn through a large number of observed points, reducing the probable error to about  $\pm 20^\circ$ . This represents an error of 4 per cent. for low values of aerial temperature and about 0.7 per cent. for high values. In the low temperature regions where the effect of the finite aerial beam is not important, application of the summation process to the resolved  $T_e$  distribution was found to reproduce the observed  $T_a$  distribution exactly, but in the high temperature regions discrepancies up to 2 per cent. were experienced. It can be seen that the combination of experimental errors and errors in analysis amount to not more than 5 per cent. in any part of the final distribution.

There is a possibility that fine details in the actual temperature distribution may not have been entirely resolved by the summation process adopted in the analysis. Such features would have to be greater than  $1^\circ$  and less than the beamwidth of the aerial ( $17^\circ$ ), otherwise they would have become apparent in the survey over the sea. It is unlikely, however, that any marked discontinuities exist in the temperature gradients, as in such cases it would not have been possible to have achieved such a high degree of consistency between the observed aerial temperatures and those computed from the resolved distribution and the aerial diagram.

The accuracy of both the distribution of aerial temperature and equivalent temperature outside the region of galactic longitudes  $140-20^\circ$  (effectively north of Declination  $+40^\circ$ ) is probably lower than that of the rest owing to the low angle of observation.





100 Mc/s. array on equatorial mounting. *Inset* : Axes in position used for survey over the sea.



## VII. ACKNOWLEDGMENTS

We wish to acknowledge the assistance of Mr. G. J. Stanley and Mr. O. B. Slee, Division of Radiophysics, C.S.I.R.O., in the construction of equipment and in the taking of observations concerned in this work.

## VIII. REFERENCES

- (1) JANSKY, K. G.—Directional studies of atmospherics at high frequencies. *Proc. Inst. Radio Engrs. N.Y.* **20** : 1920-32 (1932).
- (2) REBER, G.—Cosmic static. *Astrophys. J.* **100** : 279 (1944).
- (3) REBER, G.—Cosmic static. *Proc. Inst. Radio Engrs. N.Y.* **36** : 1215-8 (1948).
- (4) HEY, J. S., PARSONS, S. J., and PHILLIPS, J. W.—An investigation of galactic radiation in the radio spectrum. *Proc. Roy. Soc. A* **192** : 425-45 (1948).
- (5) BURGESS, R. E.—Noise in receiving aerial systems. *Proc. Phys. Soc.* **53** : 293-304 (1948).
- (6) BOLTON, J. G., and STANLEY, G. J.—Observations on the variable source of cosmic radio-frequency radiation in the constellation of Cygnus. *Aust. J. Sci. Res. A* **1** : 58-69 (1948).
- (7) BOLTON, J. G., STANLEY, G. J., and SLEE, O. B.—Positions of three discrete sources of galactic radio-frequency radiation. *Nature* **164** : 101 (1949).



# RADIO-FREQUENCY RADIATION FROM THE QUIET SUN

By S. F. SMERD\*

[*Manuscript received August 30, 1949*]

## *Summary*

The equation of transfer of radiation is used in a ray treatment of radio-frequency radiation from the solar atmosphere in the absence of solar activity. The chromosphere and the corona are represented as regions of uniform temperature. However, a range of temperatures is considered in order to allow for uncertainties in the temperature distribution in the solar atmosphere.

The intensity distribution across the solar disk is shown to depend significantly on the coronal temperature. Both limb-brightening and limb-darkening can occur but are appreciable only when the corona is optically thin, yet not transparent. A measure of the size of the radio-frequency disk is obtained in terms of the size of the optical disk.

The apparent temperature (an equivalent measure of the flux density at the earth) is found to reach a maximum as a function of frequency for each coronal temperature and as a function of coronal temperature for each frequency.

Within the limits of experimental error all observed apparent temperatures fall within the range of the theoretical values corresponding to chromospheric temperatures from  $10^4$  to  $3 \times 10^4$  °K. and coronal temperatures from  $2.5 \times 10^5$  to  $3 \times 10^6$  °K.

The effects of a possible general magnetic field of the sun are estimated in selected cases and found to be small in relation to those due to the uncertainties in the temperature of the solar atmosphere.

## I. INTRODUCTION

During recent years a great deal of observational evidence has been assembled on radio-frequency radiation from sun to earth. For a detailed discussion of the material the reader is referred to a survey paper by Pawsey(1). The results indicate that the received radiation consists of a steady component which is present irrespective of any solar activity, and variable components which are connected with sunspot and other solar activity. The former is referred to as the thermal component of the radiation or the radiation from the quiet sun. It is the object of this paper to study the properties of the radio-frequency radiation from the quiet sun. Previous theoretical treatments have been published by Ginsburg(2), Martyn(3, 4), and more recently by Waldmeier and Müller(5). The two former arrived independently at the conclusion that radiation of metre wavelengths is entirely due to the corona. In Appendix I the methods employed in these papers will be subjected to critical analysis.

In the present paper radiation is followed along the whole of any trajectory which leaves the solar atmosphere in the direction towards a distant observer. Intensity changes along the path are specified by the equation of transfer. The intensity of a ray emerging from the solar atmosphere is obtained as the integrated effect of emission, absorption, and refraction along the trajectory.

\* Division of Radiophysics, C.S.I.R.O., University Grounds, Sydney.

We then derive the following quantities associated with radiation at any one radio frequency :

- (i) The intensity distribution across the emitting disk.
- (ii) A measure of the size of the radio-frequency disk.
- (iii) The total amount of solar radiation incident upon unit area of the earth surface per unit time, i.e. the flux density at the earth.
- (iv) The apparent temperature of the sun (an alternative measure of (iii)) by which most experimental results are expressed.

In order to allow for uncertainties in our knowledge of electron temperature in the solar atmosphere, and for possible space and time variations of this temperature, the treatment covers a range of chromospheric and coronal temperatures. The chromosphere and the corona are considered as two regions of uniform temperature with a discontinuity of the temperature at the boundary. At a given frequency the temperature is determinate only over the region of origin of radiation at this frequency. Such regions differ appreciably both in extent and in position in the atmosphere at different frequencies. Thus, by associating different values of chromospheric and coronal temperatures with the regions of origin at various frequencies, we can simulate actual chromospheric and coronal temperature-gradients and any plausible transition from chromospheric to coronal temperatures.

The general treatment is carried out for zero magnetic field. The effects of a possible general magnetic field of the sun are discussed in Appendix III.

## II. THE PHYSICAL PROPERTIES OF THE SOLAR ATMOSPHERE

### (a) *General*

Current views on the relative abundance of elements in the solar atmosphere(6) lead to the conclusion that the effect of all elements other than hydrogen on the mean degree of ionization is small. These effects will be neglected and the solar atmosphere will be regarded as a fully ionized hydrogen gas.

The effective parts of the solar atmosphere for radiation of frequencies between, say,  $10^4$  and  $10$  Mc/s. are the chromosphere and the inner corona. The chromosphere is here taken to extend from 500 to 10,000 km. above the photosphere. We shall assume the solar atmosphere to be spherically symmetrical. This assumption may not always lead to a faithful representation of the actual conditions in the solar atmosphere. It is, for instance, well established from eclipse observations, that the shape of the corona changes with the solar cycle. The assumption of spherical symmetry is, however, generally employed and seems the most reasonable one to use in a simple analysis.

When the magnetic field is introduced (Appendix III) it is represented by a central dipole of moment  $8.4 \times 10^{18}$  gauss km.<sup>3</sup>(7).

### (b) *The Electron Density*

The corrected form of Baumbach's formula derived by Allen(8) is used for the electron density of the corona.\* This is

$$N = 10^8 (1.55 \rho^{-6} + 2.99 \rho^{-16}) \text{ cm.}^{-3}, \dots\dots\dots (2.1)$$

\* A similar correction has also been given by v.d. Hulst(9).

where  $\rho = R/R_0$  is the distance from the centre of the sun in terms of its optical radius  $R_0$ . The value of  $R_0$  used here is  $6.95 \times 10^5$  km.

The electron density in the chromosphere is represented by an exponential distribution

$$N = 5.7 \times 10^{11} \exp\{-7.7 \times 10^{-4}(h-500)\} \text{ cm.}^{-3}, \dots (2.2)$$

where  $h$  is the height above the photosphere in km., and

$$h = R_0(\rho - 1) \dots \dots \dots (2.3)$$

The exponential in (2.2) is that suggested by Cillié and Menzel(10), with the absolute values adjusted to join the coronal density-distribution at  $10^4$  km. above the photosphere. The electron concentration 500 km. above the base of the chromosphere is then  $3.9 \times 10^{11} \text{ cm.}^{-3}$ , compared with  $2.7 \times 10^{11} \text{ cm.}^{-3}$  as used by Cillié and Menzel, and  $5 \times 10^{11} \text{ cm.}^{-3}$  suggested by Giovanelli(11) as a mean value from available eclipse observations.

### (c) *The Electron Temperature*

In the absence of detailed observational evidence or of an established theoretical derivation, the temperature distribution in the solar atmosphere will be represented by uniform temperatures,  $T_{ch}$  and  $T_c$ , throughout the chromosphere and corona respectively. A representative range of values of  $T_{ch}$  and  $T_c$  will be used. It should be noted that a step-like temperature-distribution in conjunction with the density distribution given by (2.1) and (2.2) does not result in a continuous decrease outwards of the electron pressure, which is proportional to the product  $NT$ . The effect of a gradual increase in temperature between chromosphere and corona (compatible with, for instance, the condition of a continuous decrease outwards of the electron pressure) will be discussed in Section IV (b). The electron temperature of the corona required to interpret the observed optical spectrum is of the order of  $10^6$  °K.(12).

The electron temperature of the chromosphere deduced by Redman(13) from chromospheric line profiles is given as  $3 \times 10^4$  °K. Redman's measurements refer to an estimated height of 1500 km. and he suggests that the temperature at higher chromospheric levels would be about the same.

## III. THE TRANSFER OF RADIO-FREQUENCY RADIATION IN THE SOLAR ATMOSPHERE

### (a) *The Equation of Transfer*

The problem of the transfer of radiation in an ionized gas has been discussed by Smerd and Westfold(14). Some of the discussion is reproduced here to preserve the continuity of the argument.

We consider radiation only in the frequency interval  $f, f+df$  and refer to such radiation as  $f$ -radiation. Changes in the intensity of  $f$ -radiation along a path element  $ds$ , due to emission and absorption in an elementary cylinder (of length  $ds$  and cross-section  $dS$ ) of the medium of refractive index  $\mu$ , are related by the equation of transfer. The form of the equation used here is essentially that given by Woolley(12)

$$\frac{d(I/\mu^2)}{ds} = \frac{\eta}{\mu^2} - \kappa \frac{I}{\mu^2}, \dots \dots \dots (3.1)$$



where  $I$  is the specific intensity (subsequently abbreviated to intensity) of  $f$ -radiation in the direction of  $ds$ , i.e. the amount of  $f$ -radiation passing normally through  $dS$  into a solid angle  $d\omega$  in time  $dt$  is  $I \cdot df \cdot dS \cdot d\omega \cdot dt$ ,

$\eta$  is the volume emissivity,

$\kappa$  is the linear absorption coefficient,

$\mu$  is the refractive index of the medium.

Using the concept of optical depth introduced by Milne, we write

$$\tau = \int_0^s \kappa ds \dots\dots\dots (3.2)$$

and the equation of transfer becomes

$$\frac{d(I/\mu^2)}{d\tau} = \frac{\eta}{\mu^2\kappa} - \frac{I}{\mu^2} \dots\dots\dots (3.3)$$

Under conditions of thermodynamic equilibrium the emissivity is given by (15)

$$\eta = \mu^2 \kappa B(T), \dots\dots\dots (3.4)$$

where  $B(T)$  is the intensity of black-body radiation of frequency  $f$  and temperature  $T$ . In the case of radio frequencies this is adequately represented by the Rayleigh-Jeans radiation formula

$$B(T) = \frac{2k}{c^2} f^2 T. \dots\dots\dots (3.5)$$

Under these conditions the equation of transfer is, from (3.3), (3.4), and (3.5),

$$\frac{d(I/\mu^2)}{d\tau} = \frac{2k}{c^2} f^2 T - I/\mu^2 \dots\dots\dots (3.6)$$

The solution of this can be written as

$$\frac{Ie^\tau}{\mu^2} = \frac{2kf^2}{c^2} \int_{\tau_0}^{\tau} T e^\tau d\tau + \frac{I_0 e^{\tau_0}}{\mu_0^2}, \dots\dots\dots (3.7)$$

where the intensity is  $I_0$  and the refractive index  $\mu_0$  at an optical depth  $\tau_0$ .

For a uniform temperature region, (3.7) gives

$$I = \frac{2k}{c^2} \mu^2 f^2 T (1 - e^{\tau_0 - \tau}) + I_0 (\mu/\mu_0)^2 e^{\tau_0 - \tau} \dots\dots\dots (3.8)$$

The distribution of electron velocities in the solar atmosphere is assumed to be Maxwellian and brought about by collisions (12). It has been shown (14) that under such conditions the radio-frequency emissivity is given by (3.4). Thus equation (3.6) is the appropriate form of the equation of transfer of thermal radio-frequency radiation in the solar atmosphere.

It follows from this discussion that to evaluate the intensity of  $f$ -radiation emerging from the solar atmosphere we require:

- (i) The ray trajectories.
- (ii) The optical depth of the medium along any trajectory.
- (iii) The temperature distribution in the solar atmosphere.

(b) *The Ray Trajectories*

We now consider the paths through the solar atmosphere of those radio-frequency rays which can be received at the earth. These rays will emerge from the solar atmosphere parallel to and at a distance  $d$ , say, from the sun-earth line. We specify a point on the trajectory of such a ray by the polar coordinates  $(\rho, \theta)$  in the appropriate plane containing the sun-earth line with the pole at the centre of the sun and the line  $\theta=0$  directed towards the earth (see Fig. 1). Both the quantities  $\rho$  and  $d$  are expressed in units of the sun's optical radius,  $R_0$ .

Now the refractive index,  $\mu$ , of an ionized medium decreases with increasing electron density. It follows that a ray passing through the solar atmosphere experiences continuous bending by refraction. If the propagation of radiant

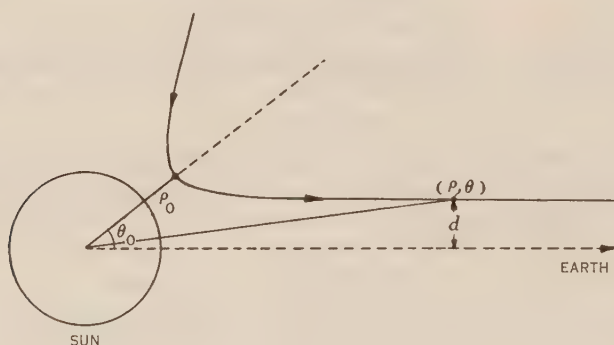


Fig. 1.—A typical ray trajectory from sun to earth.

energy is in the direction of decreasing  $\mu$ , i.e. towards the sun, the bending will be *away from* the normal to a sphere of constant  $\mu$ . A point is reached where the direction of propagation is tangential to such a sphere after which propagation will necessarily be in the direction of increasing  $\mu$ . The point where the direction of propagation changes from that of decreasing  $\mu$  to that of increasing  $\mu$  is often referred to as the “reflection point” or better as the turning point.

Westfold (unpublished data) gives the following expressions which determine a trajectory of parameter  $d$ :

$$\theta = \int_0^u \frac{du}{\sqrt{\mu^2/d^2 - u^2}}, \dots\dots\dots (3.9)$$

where  $u=1/\rho$ .

At the turning point,  $(\rho_0, \theta_0)$ ,

$$u_0 = \mu_0/d, \dots\dots\dots (3.10)$$

where  $\mu_0$  is the refractive index at  $\rho=\rho_0$ .

It follows from (3.9) that the path element,  $ds$ , of a trajectory,  $d$ , is given by

$$ds = \frac{R_0 d \rho}{\sqrt{1 - d^2/\mu^2 \rho^2}}, \dots\dots\dots (3.11)$$

since in polar coordinates

$$ds = \sqrt{(dR)^2 + R^2(d\theta)^2}.$$

The trajectories are symmetrical about the radius vector through the turning point.

In subsequent discussion the parts of the medium where  $\mu \neq 1$  at the frequency considered will be referred to as the deviating region and those parts where  $\mu = 1$  as the non-deviating region, in accordance with the terminology of ionospheric work.

### (c) The Optical Depth

The optical depth is defined by (3.2). It is convenient to measure the path length,  $s$ , from the point at which the earth-bound arm of the trajectory emerges from the solar atmosphere and in the direction *away from* the earth. Since  $\rho \rightarrow \infty$  as  $s \rightarrow 0$ , the optical depth of the point  $(\rho, \theta)$  on a trajectory is, from (3.2) and (3.11),

$$\tau(d) = \int_{\rho}^{\infty} \frac{\kappa R_0 d\rho}{\sqrt{1 - d^2/\mu^2 \rho^2}} \dots\dots\dots (3.12)$$

The difference in optical depth of two points on a trajectory will be referred to as the optical thickness between the two points.

The absorption coefficient and the refractive index are taken from the Lorentz theory. They are given by

$$\left. \begin{aligned} \kappa &= \frac{\nu x}{c\sqrt{1-x}}, \\ \mu &= \sqrt{1-x}, \end{aligned} \right\} \dots\dots\dots (3.13)$$

where  $\nu$  is the number of elastic collisions with ions per electron per unit time,  
 $c$  is the free-space velocity of light,  
 $x$  is a dimensionless quantity given by

$$x = \frac{e^2 N}{\pi m f^2},$$

in which  $e$  is the electronic charge in e.s.u. and  $m$  is the mass of the electron.

It has been shown by Smerd and Westfold(14) that there is an additional factor 4/3 in the above expression for  $\kappa$  if the free-path treatment of the Lorentz theory is replaced by a velocity-distribution treatment. From that reference

$$\nu = e^4 \left\{ \frac{\pi}{2m(kT)^3} \right\}^{\frac{1}{2}} N A_1(2), \dots\dots\dots (3.14)$$

where  $k$  is Boltzmann's constant and

$$A_1(2) = \ln\{1 + (4kT/e^2 N^{1/3})^2\} \dots\dots\dots (3.15)$$

The logarithmic term is a slowly varying function of  $N$  and  $T$  and can be taken to be constant over large parts of the solar atmosphere; a mean value of 21 for the chromosphere and 31.5 for the corona has been used in numerical work.

Substituting for  $\nu$  the absorption coefficient is, from (3.13) and (3.14),

$$\kappa = \frac{e^6 A_1(2) N^2}{\{2\pi(mkT)^3\}^{\frac{1}{2}} c f^2 \mu} \dots\dots\dots (3.16)$$



This can be written as

$$\kappa = \frac{AN^2A_1(2)}{f^2\mu T^{3/2}}, \dots\dots\dots (3.17)$$

where

$$A = \frac{e^6}{e\{2\pi(mk)^3\}^{\frac{1}{2}}}.$$

Thus the optical thickness between the points  $\rho_1, \rho_2$  on a trajectory,  $d$ , is, from (3.12) and (3.17)

$$\tau_{1,2}(d) = \frac{AR_0}{f^2} \int_{\rho_1}^{\rho_2} \frac{A_1(2)N^2}{T^{3/2}\sqrt{\mu^2 - d^2/\rho^2}} d\rho \dots\dots\dots (3.18)$$

and for a uniform temperature region with  $A_1(2)$  constant

$$\tau_{1,2}(d) = \frac{AR_0A_1(2)}{f^2T^{3/2}} \int_{\rho_1}^{\rho_2} \frac{N^2}{\sqrt{\mu^2 - d^2/\rho^2}} d\rho \dots\dots\dots (3.19)$$

Where the trajectory is parallel to the line  $\theta=0$  (in which case  $\mu=1$ ), it is clear from Figure 1 that

$$\cos \theta = \sqrt{1 - d^2/\rho^2}.$$

If, also,  $\cos \theta$  can be taken as constant between the points  $\rho_1$  and  $\rho_2$  in this non-deviating region, we can write

$$\tau_{1,2}(d) = \tau_{1,2}(0)/\cos \theta \dots\dots\dots (3.20)$$

In the case of the chromosphere where  $N=N_0e^{-\alpha h}$  (see equation (2.2)) the optical thickness in the non-deviating region along the central ray is, from (3.19) and (2.3),

$$\tau_{1,2}(0) = \frac{AN_0^2A_1(2)}{2\alpha f^2T^{3/2}} \left[ e^{-2\alpha h} \right]_{h_2}^{h_1} \dots\dots\dots (3.21)$$

Now the variation in  $\rho$  throughout the chromosphere is small ( $1.00 < \rho < 1.02$ ), so that we can, in this instance, employ (3.20) to determine  $\tau_{1,2}(d)$ .

In the deviating region the height in the chromosphere is conveniently expressed by the parameter  $x$ . The solution of (3.19) for the central ray is then

$$\tau_{1,2}(0) = \frac{2Af^2A_1(2)}{3\alpha T^{3/2}} \left( \frac{\pi m}{e^2} \right)^2 \left[ (1-x)^{\frac{1}{2}}(2+x) \right]_{x_1}^{x_2} \dots\dots\dots (3.22)$$

In the case of the corona where the density is given as a sum of power terms in  $\rho$  as in equation (2.1), the optical thickness in the non-deviating region is, from (3.19),

$$\tau_{1,2}(d) = \frac{AR_0A_1(2)10^{16}}{f^2T^{3/2}} \int_{\rho_1}^{\rho_2} \frac{2.4\rho^{-12} + 9.27\rho^{-21} + 8.94\rho^{-32}}{\sqrt{1 - d^2/\rho^2}} d\rho.$$

The general solution of this integral is cumbersome. For the central ray it reduces to

$$\tau_{1,2}(0) = \frac{AR_0A_1(2)10^{16}}{f^2T^{3/2}} \left[ 0.22\rho^{-11} + 0.44\rho^{-21} + 0.29\rho^{-31} \right]_{\rho_2}^{\rho_1} \dots\dots\dots (3.23)$$

In the deviating region, coronal  $\tau$ -values have to be obtained by numerical integration.

The emergent intensity is directly related to the function  $1 - e^{-\tau}$  (see equation (3.8)). The variation of this function with height in the solar atmosphere is shown in Figure 2 for the central ray at various frequencies.

Equation (3.7) together with the physical properties of the solar atmosphere as outlined in Section II, and the ray trajectories and optical depth as discussed in Sections III (b) and III (c), enable us to find the intensity of radio-frequency rays emerging from the solar atmosphere.

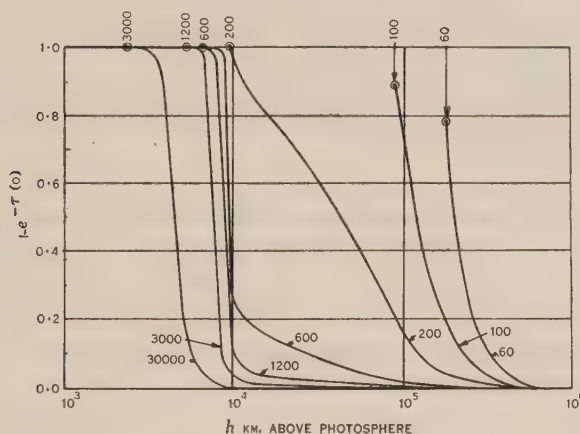


Fig. 2.—The variation of  $1 - e^{-\tau}$  with height above the photosphere for central rays of various radio frequencies;  $\tau(0)$  is the optical depth of the solar atmosphere along the central rays.

The numbers on the curves refer to the frequency in Mc/s. of the rays. Points marked  $\odot$  are the turning points at the indicated frequencies. (The turning point of the central ray of 30,000 Mc/s. is at less than  $10^3$  km. above the photosphere, and is therefore not marked in the figure.)

#### IV. RADIO-FREQUENCY RADIATION FROM SUN TO EARTH

##### (a) *The Relative Importance of the Chromosphere and Corona*

In accordance with the temperature distribution outlined in Section II, we deal with the chromosphere and the corona as two regions of uniform temperature. The assumed boundary between chromosphere and corona is at  $10^4$  km. above the photosphere. The effects of a gradual increase from chromospheric to coronal temperatures and of different values assigned to these temperatures will be discussed in Section IV (b).

In all cases in which the trajectory penetrates about 100 km. or more into the chromosphere the optical thickness of the chromosphere along the trajectory,  $\tau_{ch}$ , is such that

$$e^{-\tau_{ch}} < 1.$$

The emergent intensity along such a trajectory is then, from (3.8),

$$I_e(d) = B(T_c)(1 - e^{-\tau_c(d)}) + B(T_{ch})e^{-\tau_c(d)}, \dots \dots \dots (4.1)$$

where the subscripts "ch" and "c" refer to the chromosphere and corona respectively. The coronal and chromospheric contributions to the emergent

intensity are given by the first and second terms of the right-hand side of (4.1) respectively. When  $\tau_c \ll 1$ , the ratio,  $r$ , of these contributions is given approximately by

$$r = (T_c/T_{ch})\tau_c(d). \quad (4.2)$$

Thus, when  $T_c/T_{ch} \gg 1$ , the corona can make significant contributions to the emergent intensity of a ray even at those high frequencies at which  $\tau_c \ll 1$ .

For instance, with the electron temperatures  $T_{ch} = 3 \times 10^4$ ,  $T_c = 10^6$  °K. the ratio  $r$  is unity when  $\tau_c(d) = 0.03$ ; this value of  $\tau_c$  occurs for the central ray,  $d=0$ , at a frequency of 1430 Mc/s. (21 cm.) and for the trajectory  $d=1$  at a frequency of 2640 Mc/s. (11.4 cm.).

Table 1 gives the ratio  $r$  for various trajectories and frequencies for the above values of  $T_{ch}$  and  $T_c$ .

TABLE 1  
RATIO OF THE CORONAL TO THE CHROMOSPHERIC CONTRIBUTION TO THE EMERGENT INTENSITY OF  
VARIOUS RADIO-FREQUENCY RAYS

Frequency (Mc/s.)	Wavelength (cm.)	Trajectory ( $d$ )	$\tau_c(d)$	Ratio of Coronal to Chromospheric Contribution ( $r$ )
3000	10	0	0.007	0.23
		1	0.023	0.77
1200	25	0	0.043	1.43
		1	0.154	5.53
600	50	0	0.170	6.20
		1	0.800	$\infty$ (Turning point is in corona)

Frequencies at which the corona is practically transparent, i.e. those greater than about 10,000 Mc/s., will subsequently be referred to as *chromospheric frequencies*. Frequencies whose trajectories cannot penetrate into the chromosphere at all, i.e. those less than about 200 Mc/s., will be referred to as *coronal frequencies*. Frequencies at which both the chromosphere and the corona contribute appreciably to the emergent intensity (such as those given in Table 1) will be referred to as *intermediate frequencies*.

(b) *The Effect of the Temperature Distribution on the Radiation*

There is a large degree of uncertainty in the present knowledge of the temperature distribution in the solar atmosphere. All that can be inferred is :

(i) That the temperature is of the order  $10^4$  °K. at chromospheric heights below about  $6 \times 10^3$  km. and of the order  $10^6$  °K. over a range of coronal heights above about  $2 \times 10^4$  km.



(ii) That somewhere in this region the transition from chromospheric to coronal temperatures takes place.

Also, there is spectroscopic evidence that the coronal temperature shows spatial and temporal variations. In the absence of more precise information we represent the temperature distribution in the solar atmosphere by a two-level scheme of a uniform chromospheric and coronal temperature. We now investigate the effects on the radiation of changes in the assumed temperature distribution, taking the uncertainties into account.

In order to adapt the schematic temperature distribution for general use we have to establish the meaning of the temperature in the solar atmosphere in connexion with radio-frequency radiation. Now at a given frequency the physical properties of the solar atmosphere are determinate only over that region which is neither completely transparent, nor completely opaque, nor beyond the locus of the turning points. Such regions, referred to as "regions of origin" of radiation, will be discussed in more detail later in this section. They are found to differ appreciably both in extent and in position in the solar atmosphere at different radio frequencies. A value of the chromospheric or coronal temperature used in connexion with the properties of radiation at a given frequency is then clearly a mean electron temperature of the chromospheric or coronal region of origin at this frequency. It follows that :

(i) By assigning different values to the temperature of the region of origin at one frequency we can allow for the uncertainty in the knowledge of the temperature in this region and for the effects of possible variations of this temperature with time.

(ii) By assigning different values to the temperatures of the regions of origin at different frequencies we can simulate the effects of temperature gradients in the appropriate parts of the solar atmosphere.

In order to subject the two-level schematic representation to a critical test we consider a hypothetical atmosphere in which the transition from chromospheric to coronal temperatures is given by a simple analytical function, and compare the results given by the formal treatment with those given by the schematic method used throughout this paper. Such a comparison has been carried out for a transition of exponential form. The analysis is lengthy and has been omitted here ; its quantitative results justify the use of the two-level scheme when applied in accordance with (ii) above.

Appendix II gives a method of evaluating the intensity of radiation emerging from the solar atmosphere for any set of chromospheric and coronal temperatures from a knowledge of the intensity for one set of these temperatures. This method has been used extensively for the numerical work of this paper.

### *(c) The Variation of the Intensity across the Radio-Frequency Disk*

We now proceed to investigate the distribution of the emergent intensity across the solar disk at various frequencies and for different values of coronal and chromospheric temperatures.

It is convenient to express the emergent intensity,  $I_e$ , in terms of a temperature,  $T_e$ , given by

$$I_e = B(T_e). \quad \dots\dots\dots (4.3)$$

Following the nomenclature introduced by Pawsey(1)  $T_e$  will be referred to as the "effective temperature".

The trajectories at a frequency which can penetrate into the chromosphere are sketched in Figure 3. The outer boundary of the corona is defined by the points of zero intensity on the trajectories. The refractive index is unity at this boundary.

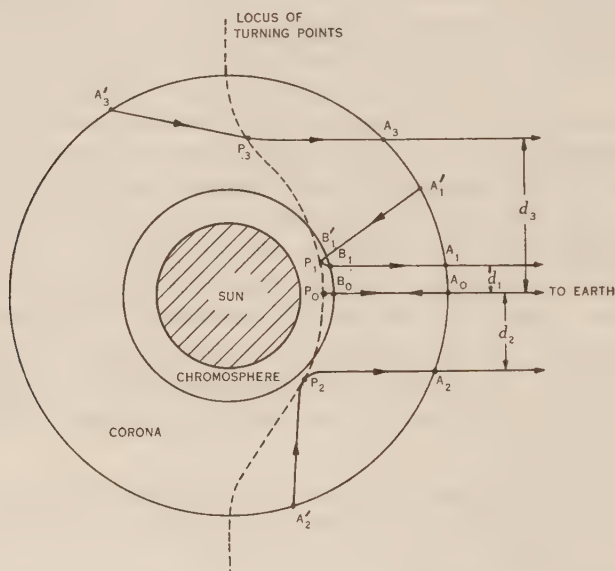


Fig. 3.—Ray trajectories at an "intermediate" frequency.

When the turning point of a trajectory is in the chromosphere (e.g.  $A_0P_0A_0$ ,  $A'_1P_1A_1$  of Fig. 3) the effective temperature of the ray is given by (4.1), (4.3), and (3.5):

$$T_e(d) = T_c(1 - e^{-\tau_c(d)}) + T_{ch}e^{-\tau_c(d)} \quad \dots\dots\dots (4.4)$$

When the complete trajectory is confined to the corona (e.g.  $A'_2P_2A_2$ ,  $A'_3P_3A_3$  of Fig. 3), the effective temperature of the ray is given by (3.8), (4.3), and (3.5):

$$T_e(d) = T_c(1 - e^{-2\tau_c(d)}), \quad \dots\dots\dots (4.5)$$

where  $\tau_c(d)$  is the optical depth of the turning point along the trajectory.

The refractive index does not appear explicitly in either (4.4) or (4.5); it is, however, contained in  $\tau$ , through both  $n$  and  $ds$ . We note that, as  $\tau_c$  ranges from 0 to  $\infty$ , the effective temperature increases continuously from  $T_{ch}$ , in the case of (4.4), and from zero, in the case of (4.5), to  $T_c$ . As the distance,  $d$ , of the trajectory from the centre of the disk increases, the optical thickness,  $\tau_c(d)$ , (and with it the effective temperature  $T_e(d)$ ) has a tendency (a) to increase

due to the lengthening of the path between the same two heights in the atmosphere, and (b) to decrease due to the increasing height of the turning point.

It follows that in the case of (4.4) the effective temperature must increase towards the limb; in the case of (4.5) either limb-brightening or limb-darkening can occur. In all cases the effective temperature tends to zero for sufficiently large values (always greater than unity) of  $d$ . Both limb-brightening and limb-darkening are significant only when  $(1 - e^{-\tau_c})$  varies appreciably with  $\tau_c$ , i.e. when the corona is optically thin yet not transparent.

The variation of the effective temperature across the disk is shown for various frequencies in Figure 4. Figure 5 gives the variation of the effective temperature with frequency for different distances from the centre of the disk. Both figures refer to the electron temperatures  $T_{ch} = 3 \times 10^4$ ,  $T_c = 10^6$  °K. Figure 6 shows the variation of the effective temperature across the disk for different coronal temperatures at selected frequencies.

Certain general conclusions can be drawn from these figures.

*At Chromospheric Frequencies (see Fig. 4)*

- (i) The effective temperature of any ray within the optical disk ( $0 \leq d \leq 1$ ) is just the chromospheric temperature, since the medium is effectively of infinite optical depth along such trajectories.
- (ii) For rays outside the optical disk ( $d > 1$ ) the optical thickness of the medium, and with it the effective temperature, falls off rapidly as the parameter  $d$  increases.
- (iii) The disk at these frequencies is effectively one of uniform intensity (under the assumption of a uniform chromospheric temperature) and has a well-defined limb.
- (iv) Limb-brightening or limb-darkening across the optical disk would indicate a chromospheric temperature gradient.

*At Coronal Frequencies (see Figs. 4 and 6 (d), (e), (f))*

- (i) The optical thickness, and with it the effective temperature, decreases gradually with increasing distance from the centre of the disk.
- (ii) At the higher frequencies limb-brightening can occur within the optical disk.
- (iii) At low values of coronal temperature (e.g.  $3.3 \times 10^5$  °K.) the intensity distribution is marked by a large uniform centre. At higher values of coronal temperature the size of this uniform centre diminishes.
- (iv) The overall size of the emitting disk increases as the assumed value of the coronal temperature is decreased.

*At Intermediate Frequencies (see Figs. 4 and 6 (a), (b), (c))*

- (i) Pronounced limb-brightening can occur.
- (ii) The degree of limb-brightening at any one frequency depends markedly on the assumed value of coronal temperature.

Having evaluated the effective temperature of any ray at any one frequency for a range of chromospheric and coronal temperatures, we have essentially



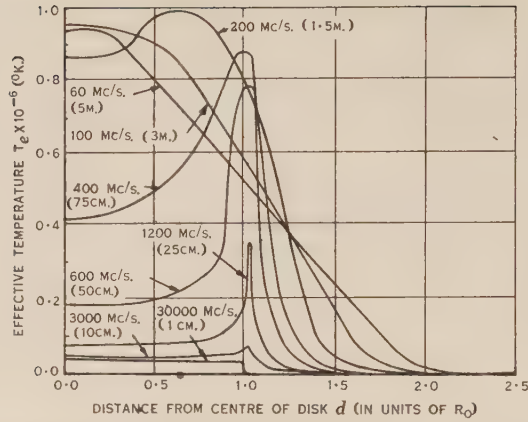


Fig. 4.—The variation of the effective temperature with distance from the centre of the disk at different radio frequencies. The values used for the chromospheric and coronal temperatures are  $3 \times 10^4$  and  $10^6$  °K. respectively.

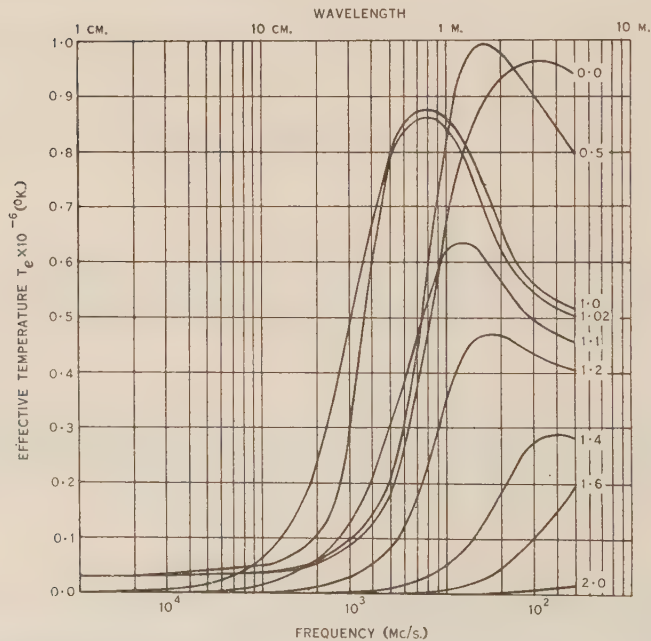


Fig. 5.—The variation of the effective temperature with frequency at different distances from the centre of the disk. The numbers on the curves refer to the distance,  $d$ , from the centre of the disk in units of  $R_0$ . The values used for the chromospheric and coronal temperatures are  $3 \times 10^4$  and  $10^6$  °K. respectively.

determined the characteristics of radiation from the quiet sun at this frequency. The remainder of this section is devoted to the interpretation of these characteristics and to the evaluation of some dependent quantities which facilitate comparison with observation.

#### (d) Regions of Origin of Radiation

For the purpose of discussing regions of origin of radiation received along a trajectory we distinguish two types of trajectories, though there is a continuous

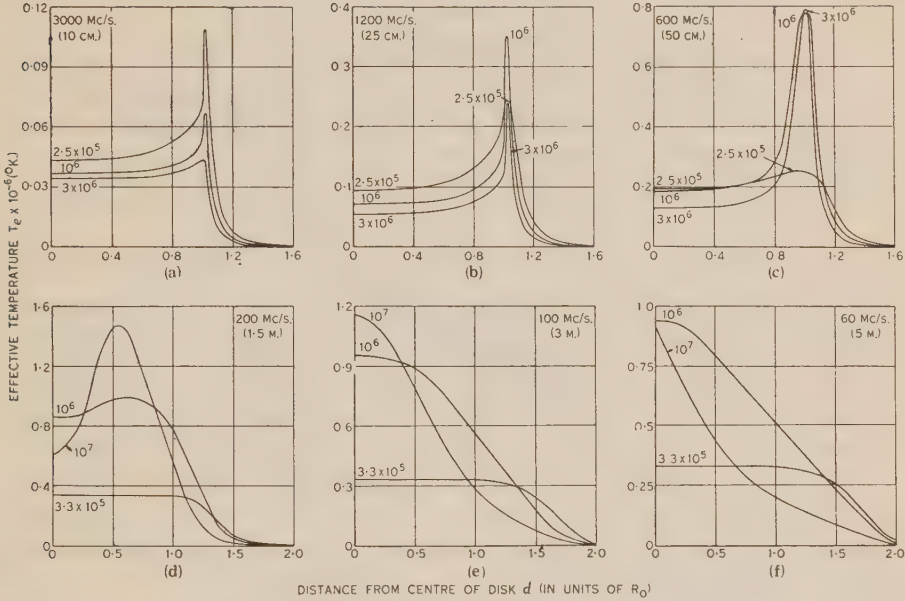


Fig. 6.—The variation of the effective temperature with distance from the centre of the disk for different coronal temperatures at selected frequencies. The numbers on the curves refer to the coronal temperature in degrees K.

transition between these. The first type comprises trajectories whose turning points are at an “effectively infinite” optical depth, defined by

$$e^{-\tau} \leq \varepsilon,$$

where  $\varepsilon < 1$ .

In quantitative work we shall use  $\varepsilon = 0.01$ . Hence the optical depth is to be considered “effectively infinite” when  $\tau \geq 4.6$ . The second type comprises the remaining trajectories, namely, those for which the optical depth is finite (i.e.  $\tau < 4.6$  in our case) at the turning point.

We can now stipulate the inner limit of the region of origin in both cases. In the first case the inner limit is the highest point in the solar atmosphere at which the optical depth along the trajectory is “effectively infinite”; in the second case this limit is just the height of the turning point.

The outer limit is more difficult to specify. Formally it lies at infinity in the case of an unbounded medium. It is, however, more useful to determine an “effective outer limit” beyond which the flow of radiation (outwards along

the trajectory) ceases to change appreciably. Thus we define the outer limit as the point along a trajectory at which the intensity,  $I$ , is related to the emergent intensity,  $I_e$ , by

$$I = (1 - \delta)I_e,$$

where  $\delta \ll 1$ .

In quantitative work we shall use the two values  $\delta = 0.1$  and  $\delta = 0.01$ .

The inner and outer limits have been calculated in this way for the central ray at selected frequencies and are given in Table 2.

TABLE 2  
 $I_e(0)$  IS THE EMERGENT INTENSITY OF THE CENTRAL RAY

Frequency (Mc/s.)	Wavelength	Inner Limit (km. above photosphere)	Outer Limit (km. above photosphere) for	
			$0.9 I_e(0)$	$0.99 I_e(0)$
30000	1 cm.	$3.0 \times 10^3$	$5.7 \times 10^3$	$8.1 \times 10^3$
3000	10 cm.	$6.0 \times 10^3$	$3.0 \times 10^4$	$7.6 \times 10^4$
1200	25 cm.	$6.8 \times 10^3$	$7.6 \times 10^4$	$1.7 \times 10^5$
600	50 cm.	$7.9 \times 10^3$	$1.0 \times 10^5$	$2.2 \times 10^5$
200	1.5 m.	$9.6 \times 10^3$	$1.3 \times 10^5$	$2.9 \times 10^5$
100	3 m.	$8.7 \times 10^4$	$2.3 \times 10^5$	$4.4 \times 10^5$
60	5 m.	$1.7 \times 10^5$	$3.3 \times 10^5$	$5.5 \times 10^5$

The heights of the inner and outer limits tend to increase with increasing distance,  $d$ , of the trajectory from the centre of the disk.

(e) *The Total Amount of Radiation from Sun to Earth*

The total amount of radiation per unit time, unit frequency interval, and unit solid angle from the sun to a distant observer is given by

$$E = 2\pi R_0^2 \int_0^\infty I_e(d) d\delta d, \dots\dots\dots (4.6)$$

where  $\delta d$  denotes the differential of  $d$ . For practical purposes it is convenient to replace the upper limit of this integral by a finite value,  $d_p$ , of  $d$  such that the value of  $E$  is not appreciably affected.

Figure 7 is a plot against frequency of  $E$  in units of watts per c/s. per unit solid angle and also of the flux density at the earth, its equivalent measure, in watts per c/s. per square metre. (The mean distance between sun and earth is taken to be  $1.5 \times 10^{11}$  m.). A typical curve showing the variation of  $E$  with  $f^2$  (i.e. according to the frequency law of black-body radiation from a source of given temperature and size) is included in this figure for comparison.

(f) *The Apparent Temperature of the Sun at Radio Frequencies*

It is convenient at this stage to introduce the concept of apparent temperature(1). The apparent temperature,  $T_a$ , is defined by attributing the



flux density at the earth to the optical disk supposedly radiating as a black body of temperature  $T_a$ . Thus

$$T_a = \frac{c^2 E}{2\pi R_0^2 k f^2} \dots\dots\dots (4.7)$$

It should be noted that apparent temperatures greater than the maximum temperature of the region of origin can be obtained since the size of the radio-frequency disk can exceed that of the optical disk.

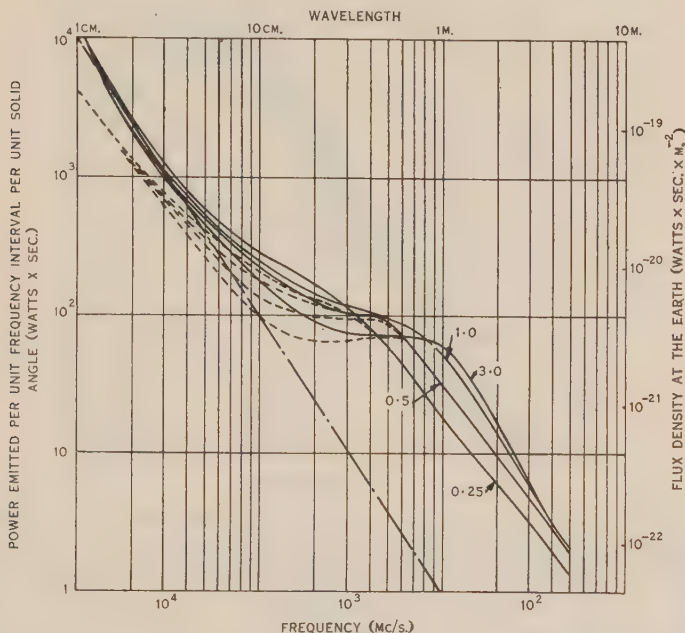


Fig. 7.—The variation of the amount of solar radiation with frequency for different chromospheric and coronal temperatures. The numbers on the curves refer to the coronal temperature in  $10^6$  °K. Two values of chromospheric temperature are used:  $T_{ch} = 3 \times 10^4$  °K. (continuous lines);  $T_{ch} = 10^4$  °K. (broken lines). Black-body radiation at any one temperature would appear in this figure as a straight line parallel to the line shown thus

Figure 8 gives the apparent temperature for a range of frequencies from 30,000 Mc/s. (1 cm.) to 60 Mc/s. (5 m.) for the electron temperatures  $T_{ch}$ :  $10^4$ ,  $3 \times 10^4$  °K., and  $T_e$ :  $2.5 \times 10^5$ ,  $5 \times 10^5$ ,  $10^6$ ,  $3 \times 10^6$  °K. Included in this figure are experimental data and a mean value of Martyn's(4) temperatures for the "ordinary" and "extraordinary" radiation, which are taken from a paper by Pawsey(1). Figure 9 shows the variation of the apparent temperature with coronal temperature in the range from  $2.5 \times 10^5$  to  $10^7$  °K. at various frequencies. The following general conclusions can be drawn from these figures :

*At Chromospheric Frequencies (see Fig. 8)*

The apparent temperature is approximately the same as the chromospheric temperature. (When, however, a temperature gradient is attributed to the chromosphere, the apparent temperature is a mean temperature of the chromospheric region of origin.)

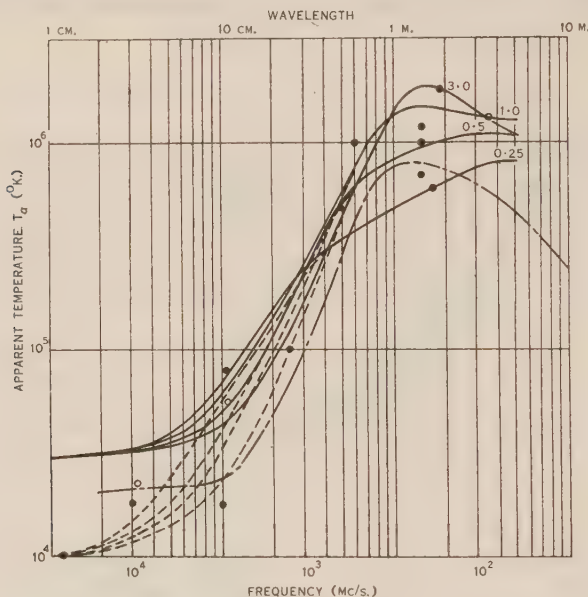


Fig. 8.—The variation of the apparent temperature with frequency for different chromospheric and coronal temperatures. The numbers on the curves refer to the coronal temperature in  $10^6$  °K. Two values of chromospheric temperature are used:  $T_{ch} = 3 \times 10^4$  °K. (continuous lines);  $T_{ch} = 10^4$  °K. (broken lines). The points and the remaining curve are taken from a paper by Pawsey(1), viz.:

- points obtained from series of observations,
- o eclipse observation,
- · — a theoretical curve due to Martyn.

*At Coronal Frequencies (see Figs. 8 and 9)*

- (i) There is a maximum of the apparent temperature as a function of frequency for each of the assumed coronal temperatures.
- (ii) The frequency at which this maximum occurs decreases as the coronal temperature falls below  $10^6$  °K.
- (iii) For coronal temperatures between  $10^6$  and  $10^7$  °K. the maximum apparent temperature occurs at or near 200 Mc/s. (1.5 m.). (The  $10^7$  °K. curve has been omitted from Fig. 8 to avoid congestion.)
- (iv) There is a maximum of the apparent temperature as a function of coronal temperature for each frequency. Thus, there are, in general, two coronal temperatures which may yield the same apparent temperature.

*At Intermediate Frequencies (see Figs. 8 and 9)*

- (i) The apparent temperature drops sharply with frequency.
- (ii) Same as (iv) above.
- (iii) The effect of the chromospheric temperature on the apparent temperature increases with frequency.

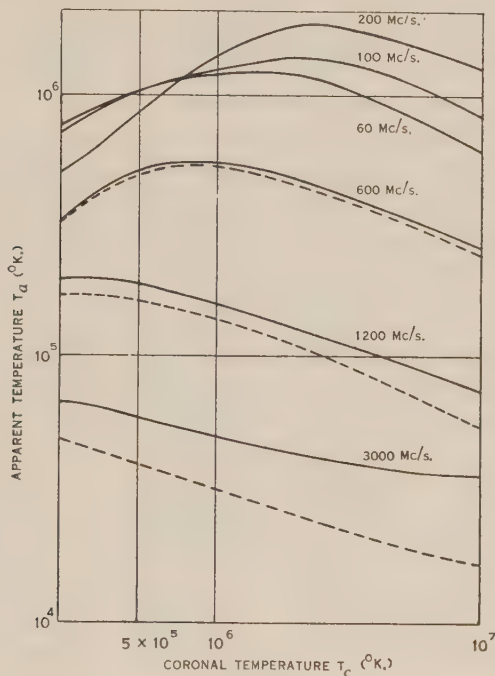


Fig. 9.—The variation of the apparent temperature with coronal temperature at selected frequencies. Two values of chromospheric temperature are used :

$T_{ch} = 3 \times 10^4$  °K. (continuous lines) ;  $T_{ch} = 10^4$  °K. (broken lines).

#### (g) The Size of the Radio-Frequency Disk

We shall specify the radius,  $d_p$ , of the effective radio-frequency disk by the value of the parameter,  $d$ , at which the effective temperature falls to a small percentage,  $p$ , of (a) the apparent temperature, and (b) the maximum effective temperature.

Figure 10 shows the variation with frequency of this radius as defined by (a) and (b) for two values of  $p$ . A coronal temperature of  $10^6$  °K. has been used.

The proportion of the total amount of radiation from beyond  $d_{10}$ , as defined by (a), increases from 0.5 per cent. at 30,000 Mc/s. (1 cm.) to 6.6 per cent. at 60 Mc/s. (5 m.) and is less than 1 per cent. at all frequencies beyond  $d_1$ .

The size of the radio-frequency disk increases with decreasing coronal temperature (see Fig. 6).



## V. COMPARISON OF THEORY AND EXPERIMENT

(a) *The Apparent Temperature*

The available radio-frequency measurements of the apparent temperature in the range of frequencies from 30,000 Mc/s. (1 cm.) to 75 Mc/s. (4 m.) have been analysed by Pawsey and Yabsley(16). There is general agreement between the observed and the theoretical apparent temperatures. Within the limits of experimental error, all observed values of  $T_a$  can be accounted for by theoretical values based on mean electron temperatures (of the regions of origin) which are consistent with temperatures obtained by optical methods.

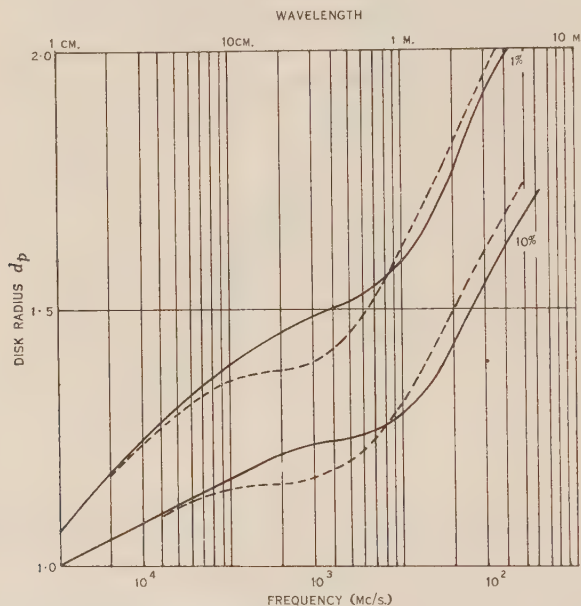


Fig. 10.—The variation of the disk radius with frequency. The radius,  $d_p$  (in units of  $R_0$ ), of the disk is specified by the ratio ( $p$  per cent.) of the effective temperature at the limb to (a) the apparent temperature (continuous lines) and (b) the maximum effective temperature (broken lines). The curves refer to a million-degree corona.

At chromospheric frequencies it has been shown that the apparent temperature is a mean electron temperature of the region of origin. Thus we interpret the observed value,  $10^4$  °K., of the apparent temperature at 30,000 Mc/s. (1 cm.) as a mean electron temperature over the range of electron densities from about  $8 \times 10^{10}$  to about  $10^{10}$  cm.<sup>3</sup>. With the density distribution of Section II this corresponds to a range in height from  $3 \times 10^3$  to  $5.7 \times 10^3$  km. above the photosphere. It seems difficult to reconcile this result with Redman's(13) value of  $3 \times 10^4$  °K. at an estimated height of  $1.5 \times 10^3$  km. Agreement could, however, be achieved by postulating a "trough" in the chromospheric temperature distribution above  $1.5 \times 10^3$  km. The two measurements could also be brought into line if the appropriate range of electron densities occurred closer to the photosphere than is suggested above. A more attractive possibility has

been suggested by Giovanelli (personal communication) in the form of a chromospheric model consisting of an atmosphere penetrated by columns at a different temperature. The apparent temperature resulting from this model at chromospheric frequencies would be a mean value of the temperatures of the two constituent parts of the medium. Giovanelli adduces reasons why optical measurements would result in temperatures at or near the higher of the two temperatures.

At coronal and intermediate frequencies we found in Section IV (f) that a given value of apparent temperature can result from two different values of coronal temperature. The lower values calculated for the observed apparent temperatures range from  $3 \times 10^5$  °K. to  $3 \times 10^6$  °K., and the higher values from  $2 \times 10^6$  °K. to  $3 \times 10^7$  °K. Current views on coronal temperatures(12) indicate that the larger values of the higher temperature range are unlikely ; however, we cannot discard the higher temperatures on apparent-temperature evidence alone. It may in certain cases be possible to resolve this ambiguity from a knowledge of either of the following radio-frequency properties :

- (i) The general form of the intensity distribution across the disk, in cases where the two coronal temperatures are widely separated (see Fig. 6).
- (ii) The gradient  $(dT_a/df)_f$ , in cases where this gradient differs appreciably at the two coronal temperatures (see Fig. 8).

TABLE 3

Observer	Frequency (Mc/s.)	Wave- length (cm.)	Method	Results	
				Intensity Distribution	Approximate Radius of Disk in Units of $R_0$
Dicke and Beringer(17) (1946)	24,000	1.25	Eclipse	Approximately uni- form	1
Sander(18) (1947) ..	9,375	3.2	Eclipse	Suggestion of limb-bright- ening	1
Covington(19) (1947)	2,800	10.7	Eclipse	Inconclusive	Inconclusive
Reber(20) (1946) ..	480	62.5	Series of measurements	—	1
Reber(21) (1948) ..	480	62.5	Not stated	—	1.4

(b) *The Intensity Distribution Across the Radio-Frequency Disk*

The published evidence concerning intensity distribution and disk sizes is summarized in the table below.

Apart from Reber's earlier figure the above results, where at all definite, are consistent with those obtained in this paper (see Figs. 4 and 10).

More substantial evidence has been obtained in recent eclipse observations by Christiansen, Yabsley, and Mills(22) at a frequency of 600 Mc/s. (50 cm.). These show that at the time of observation part of the radiant energy must have come from beyond the optical limb. When the effects of localized active areas on the solar disk are subtracted, the remaining intensity distribution is consistent with that shown for this frequency and a million-degree corona in Figures 4 and 6.

## VI. ACKNOWLEDGMENTS

The author wishes to thank Dr. J. L. Pawsey, Mr. K. C. Westfold, and Mr. J. P. Wild, Division of Radiophysics, C.S.I.R.O., for helpful criticism and advice in the preparation of this paper.

## VII. REFERENCES

- (1) PAWSEY, J. L.—*J. Instn. Elect. Engrs.* **97** (1950).
- (2) GINSBURG, V. L.—*C.R. Acad. Sci. U.R.S.S.* **52** : 487 (1946).
- (3) MARTYN, D. F.—*Nature* **158** : 632 (1946).
- (4) MARTYN, D. F.—*Proc. Roy. Soc. A* **193** : 44 (1948).
- (5) WALDMEIER, M., and MÜLLER, H.—*Astr. Mitt. Zurich* No. 155 (1948).
- (6) ALLEN, C. W., and WOOLLEY, R. v. D. R.—*Mon. Not. R. Astr. Soc.* **108** : 292 (1948).
- (7) CHAPMAN, S.—*Mon. Not. R. Astr. Soc.* **103** : 117 (1943).
- (8) ALLEN, C. W.—*Mon. Not. R. Astr. Soc.* **107** : 426 (1947).
- (9) v. D. HULST, H. C.—*Astrophys. J.* **105** : 471 (1947).
- (10) CILLÉ, G. G., and MENZEL, D. H.—*Harv. Coll. Obs. Circ.* No. 410 (1935).
- (11) GIOVANELLI, R. G.—*Aust. J. Sci. Res. A* **1** : 360 (1948).
- (12) WOOLLEY, R. v. D. R.—*Aust. J. Sci. Suppl.* **10** (2) : i (1947).
- (13) REDMAN, R. O.—*Mon. Not. R. Astr. Soc.* **102** : 140 (1942).
- (14) SMERD, S. F., and WESTFOLD, K. C.—*Phil. Mag.* **40** : 831 (1949).
- (15) MILNE, E. A.—*Handb. Astrophys.* **3** : 80 (1930).
- (16) PAWSEY, J. L., and YABSLEY, D. E.—*Aust. J. Sci. Res. A* **2** : 198 (1949).
- (17) DICKE, R. H., and BERINGER, R.—*Astrophys. J.* **103** : 375 (1946).
- (18) SANDER, K. F.—*Nature* **159** : 506 (1947).
- (19) COVINGTON, A. E.—*Nature* **159** : 405 (1947).
- (20) REBER, G.—*Nature* **158** : 945 (1946).
- (21) REBER, G.—*Proc. Inst. Radio Engrs.* **36** : 88 (1948).
- (22) CHRISTIANSEN, W. N., YABSLEY, D. E., and MILLS, B. Y.—*Aust. J. Sci. Res. A* **2** : 506 (1949).
- (23) APPLETON, E. V.—*J. Instn. Elect. Engrs.* **71** : 642 (1932).
- (24) BABCOCK, H. D.—*Publ. Astr. Soc. Pacif.* **60** : 244 (1948).

## APPENDIX I

### *Previous Treatments*

The theory of radio-frequency radiation from the quiet sun has previously been investigated by Ginsburg (1946), Martyn (1946, 1948), and Waldmeier and Müller (1948).

Ginsburg(2) obtains the amount of radiation received at the earth from the Rayleigh-Jeans formula by assuming a radio-frequency disk greater than the optical disk which radiates as a black body of coronal temperature. The radius of the assumed disk at a given frequency is stated to be "a certain average value of the ratio  $r/r_0$  corresponding to the region responsible for the radiation". (Here  $r/r_0$  measures the distance of points from the centre of the sun in terms of the sun's radius.) The use of the coronal temperature as a mean effective temperature over a vaguely defined radio-frequency disk does not lend itself to derive the flux density of radiation at the earth to more than an order of magnitude.

Martyn(4) predicts the effective temperature of a ray by the use of two criteria. The first of these states that if the optical depth of the turning points is greater than or equal to unity, the effective temperature of the ray is given by the temperature at the point  $\tau=1$  along the trajectory. The reference given for this criterion gives the average optical depth  $\tau=1$  traversed by radiation emerging from a semi-infinite, uniform temperature medium. There is, therefore, no necessary connexion between  $\bar{\tau}$  and the optical depth  $\tau_e$ , at which the temperature is equal to the effective temperature of the ray. Martyn's second criterion states that if the optical depth,  $\tau_p$ , of the turning point is less than unity, the effective temperature of a ray is given by

$$T_e = T_p(1 - e^{-2\tau_p}),$$

where  $T_p$  is the temperature at the turning point. This criterion is ascribed to Kirchhoff's law. It has subsequently been substantiated by Woolley(12) from the equation of transfer and shown to yield the correct answer for the effective temperature of a ray emerging from a uniform temperature medium with zero incident radiation and unit refractive index at incidence and emergence.

Martyn finds a mean value,  $T_e$ , of the effective temperature by weighting any annulus of the disk proportional to its projected area. If we denote the radius of the radio-frequency disk by  $R_{of}$  (as distinct from the optical radius  $R_0$ ) the function

$$E = \pi R_{of}^2 B(\bar{T}_e)$$

represents the amount of radiation from sun to earth per unit time, unit frequency interval, and unit solid angle. In terms of the apparent temperature,  $T_a$ ,

$$E = \pi R_{of}^2 B(T_a).$$

Thus, the ratio of the apparent temperature to a mean effective temperature is

$$T_a/\bar{T}_e = R_{of}^2/R_0^2.$$

Martyn does not evaluate  $R_{of}$  and hence leaves the apparent temperature indefinite to the extent given by the ratio  $R_{of}^2/R_0^2$ . Thus Martyn's values  $\bar{T}_e$  can be compared with experimental values of  $T_a$  only after making an arbitrary assumption concerning  $R_{of}$ , such as for instance the assumption  $R_{of} = R_0$  implied by Pawsey(1).

Where the general magnetic field is introduced an approximate treatment is used to compute the extraordinary radiation from the previously obtained radiation in the field-free case. Otherwise the same two criteria are used to



predict the effective temperature and the same comments as to their relevance and validity apply.

Waldmeier and Müller(5) derive the absorption coefficient from Kramers's formula for free-free transitions. Apart from differences in the logarithmic term, their absorption coefficient is independent of the refractive index. The effect of the refractivity of the medium on the ray trajectories is also neglected. The trajectories are assumed to be rectilinear and not to reach a turning point. The resultant error in the effective temperature is negligibly small for all rays at the highest frequencies considered and for the outer rays at all frequencies. For the other rays the error increases considerably with decreasing frequency. For instance, the apparent temperature at 60 Mc/s. (5 m.) under the specified conditions of electron temperature and density is about three times too large.

## APPENDIX II

### *The Variation of Intensity with the Temperature of the Region of Origin*

Consider a region of known electron-density distribution, to which we ascribe the temperatures  $T_1$  and  $T_2$  in turn. Let  $\tau_1$  and  $\tau_2$  be the corresponding optical thicknesses along a path through this region (see inset to Fig. 11). Then it follows from (3.19), assuming  $A_1(2)$  constant, that

$$\tau_2 = \tau_1 / \alpha^{3/2}, \dots\dots\dots (11.1)$$

where  $\alpha = T_2/T_1$ .

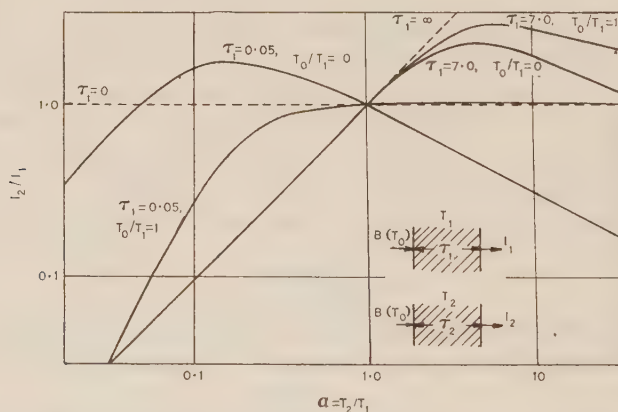


Fig. 11.—The curves show the ratio,  $I_2/I_1$ , of intensities associated with two temperatures,  $T_1$  and  $T_2$ , of the region of origin plotted against  $T_2/T_1$  ( $=\alpha$ ) for two values of  $\tau_1$ , the optical thickness along a path through this region at the temperature  $T_1$ . The incident intensity is written as  $B(T_0)$ .

If we denote the incident intensity by  $B(T_0)$ , the ratio of the emergent intensities corresponding to the temperatures  $T_2$  and  $T_1$  is

$$\frac{I_2}{I_1} = \frac{\alpha - e^{-\tau_1/\alpha^{3/2}}(\alpha - T_0/T_1)}{1 - e^{-\tau_1}(1 - T_0/T_1)} \dots\dots\dots (11.2)$$

Figure 11 shows  $I_2/I_1$  versus  $T_2/T_1$  ( $=\alpha$ ) for a large and a small value of  $\tau_1$  (7 and 0.05 respectively), with  $T_0/T_1=0$  and  $T_0/T_1=1$  in each case.

With  $\tau_1 = \infty$ , the emergent intensity increases with the temperature of the medium. With  $\tau_1 = 0$ , the emergent intensity is  $B(T_0)$  for all values of  $\alpha$ . For finite values,  $0 < \tau_1 < \infty$ , there is a maximum value of  $I_2/I_1$  as a function of  $\alpha$  for every  $\tau_1$ . Physically this means that the emergent intensity does not necessarily increase with the temperature of the medium—for at the same time the medium becomes optically “thinner”.

Having derived a value,  $I_1$ , of the intensity of a ray emerging from the solar atmosphere assuming one set of chromospheric and coronal temperatures, we can use equation (11.2) to evaluate readily the value,  $I_2$ , assuming any other set of temperature values.

### APPENDIX III

#### *The Effects of the General Magnetic Field of the Sun*

In accordance with the results of magneto-ionic theory (Appleton 23) there are two possible modes of propagation in the presence of a magnetic field. These are conventionally referred to as the ordinary and the extraordinary modes and will be denoted by “ $o$ ” and “ $e$ ” respectively. If the dimensionless quantity  $y = f_H/f$  (where  $f_H = \frac{eH}{2\pi mc}$  is the gyro-frequency) is such that  $y \ll 1$ ,

propagation in both modes is essentially that of the field-free case. With the sun’s general magnetic field represented by a central dipole of moment  $8.4 \times 10^{18}$  gauss km.<sup>3</sup> this condition is fulfilled at the chromospheric frequencies.

At intermediate and coronal frequencies the values of  $y$  along most trajectories are in the range  $0 < y < 1$ .

The dimensionless quantity  $z = \nu/2\pi f$  (where  $\nu$  is the collision frequency) is such that  $z \ll 1$  at all frequencies and all parts of the medium.

Radio-frequency propagation under such conditions has been discussed by Westfold (unpublished data), who gives the following formulae appropriate for small values  $y < 1$ :

$$\begin{aligned}\tau_o &= \tau - \tau', \\ \tau_e &= \tau + \tau', \\ \tau' &= 2 \int_s^\infty \kappa y |\cos \theta| ds,\end{aligned}$$

where  $\theta$  is the angle between the magnetic field,  $H$ , and the direction of propagation, and  $\kappa$  and  $\tau$  are the absorption coefficient and the optical depth in the absence of a magnetic field.

We denote the effective temperature of a ray in the absence of a magnetic field by  $T_e$ , and that of the “ $o$ ” and “ $e$ ” components by  $T_{e,o}$  and  $T_{e,e}$  respectively. Writing  $T_{e,o} + T_{e,e} = T_{e,oe}$ , the percentage change in the effective temperature due to the magnetic field is given by

$$\frac{100(T_{e,oe} - T_e)}{T_e},$$

and the percentage polarization is given by

$$p = \frac{100 |T_{e,e} - T_{e,o}|}{T_{e,oe}}.$$

The limiting polarization (i.e. the polarization at emergence from the solar atmosphere) is found to be circular along most trajectories, the "o" and "e" components being of opposite sense. For trajectories close to the equator and near the limb of the radio-frequency disk the limiting polarization is linear.

If the receiving device consists of two aerials which accept the two circularly-polarized components of radiation, the aerial temperatures will be  $T_{e,o}$  and  $T_{e,e}$  respectively in the first case, and both equal to  $\frac{1}{2}T_{e,o,e}$  in the second case.

Table 4 gives the upper limit of magnetic-field effects (corresponding to  $(y | \cos \theta |)_{max}$ ) at intermediate frequencies and for a range of  $\tau$ -values appropriate to chromospheric and coronal trajectories.

TABLE 4

Frequency (Mc/s.)	Wavelength (cm.)	Type of Trajectory	$\tau_c$	Upper Limit of :	
				Percentage Change in Effective Temperature	Percentage Polar- ization
600	50	Chromospheric ..	0.2	-1.6	35
		Coronal .. ..	0.3	-4.0	35
		Coronal .. ..	0.8	-6.5	21
3000	10	Chromospheric ..	0.007	0	1.6
		Chromospheric ..	0.02	0	3.8
		Coronal .. ..	0.035	-0.15	8.6

At coronal frequencies the approximations for small values  $y < 1$  can be applied to the outer trajectories only.

A numerical example at 100 Mc/s. is given in Table 5.

TABLE 5

Trajectory (d)	Percentage Change In Effective Temperature	Percentage Polarization
0.0	-1.3	1.3
0.5	-1.7	1.8
1.0	1.2	30
	(Upper limit)	
1.6	1.6	59
2.0	0	36

The magnetic-field effects at other coronal frequencies are likely to be of the same order of magnitude.

From the above examples we conclude that the effects of a general magnetic field of the sun on the effective and the apparent temperatures of radio-frequency radiation are smaller than those conditioned by the uncertainty in the knowledge of electron temperatures and densities. Moreover, some doubt has recently been thrown(24) on the constancy, or even the existence, of the sun's general magnetic field. Hence the field-free treatment can be taken as representative of radio-frequency radiation from the quiet sun.

However, differential-polarization experiments may disclose the presence and magnitude of a general magnetic field of the sun.



# SOLAR RADIATION AT A WAVELENGTH OF 3.18 CENTIMETRES

By H. C. MINNETT\* and N. R. LABRUM\*

[Manuscript received August 31, 1949]

## Summary

Solar radiation at a wavelength of 3.18 cm. has been measured over a period of three months. The received intensity was found to vary from day to day and the changes are shown to be closely associated with sunspots. The equivalent black-body temperature of the sun over this period, in the absence of sunspots, was 19,300 °K., with a probable error of  $\pm 7$  per cent. The temperature increased by 8 °K. per unit increase of sunspot area (one unit equals  $10^{-6}$  times the area of the sun's visible disk). This increase is much less than that at longer microwavelengths.

Sudden increases of radiation at 3.18 cm., caused by disturbed conditions in the sun, were found to be rare. A number of bursts were observed and a comparison is made with records of longer wave solar radiation and other phenomena of solar origin.

Observations were made during the solar eclipse of November 1, 1948 and the results are consistent with either of two simple brightness distributions on the sun's disk. In the first of these, 74 per cent. of the energy is emitted uniformly by the sun's visible disk and the remaining 26 per cent. by a bright ring around the circumference; in the second, the whole of the radiation comes from a uniform disk of diameter 1.1 times that of the visible sun.

## I. INTRODUCTION

Although Southworth's(1) early measurements of radio-frequency radiation from the sun were at centimetre wavelengths, observations by later workers were made mainly at wavelengths between 1 and 5 metres. This was due to interest in the discovery that the intensity of radiation from the undisturbed or "quiet" sun at the longer wavelengths is very much greater than the black-body value(2, 3). The explanation of this result was given by the theoretical work of Ginsburg(4) and Martyn(5, 6) which predicts the level in the solar atmosphere at which radiation of given wavelength originates. The theory shows that the chromosphere is the source of most of the centimetre and decimetre radiation so that measurements at these wavelengths should lead to a better understanding of this region of the sun's atmosphere.

In particular, two quantities which can be measured and which are of interest in the theory of the "quiet" sun are the equivalent temperature of the disk and the distribution of brightness across it. Theoretically the latter could be determined by scanning the disk with a narrow aerial beam; in practice it is not feasible to produce a sufficiently narrow beam and an interference technique must be used. Alternatively, the distribution can be deduced from observations of the manner in which the received intensity varies during an eclipse of the sun.

\* Division of Radiophysics, C.S.I.R.O., University Grounds, Sydney.

In addition to the relatively constant radiation associated with the "quiet" sun, there is a variable component of radiation which is associated with sunspots. Sudden increases of intensity or "bursts" are also found to occur. The cause of these disturbed conditions is not yet understood and microwave observations are therefore of some interest. In measurements of this type it is necessary to examine the changes in intensity over an extended period of time.

Following Southworth's experiments at 1, 3, and 10 cm., the factors discussed above were studied by several workers at a number of centimetre and decimetre wavelengths. Measurements were made at 1.25 cm. by Dicke and Beringer(7) and at 3.2 cm. by Sander(8). More extensive observations by Reber(9) (62.5 cm.), Covington(10) (10.7 cm.), and Lehany and Yabsley(11) (25 and 50 cm.) showed that the radiation at these wavelengths contained a component which varied in a manner correlating with the sunspot area. Bursts of radiation were also detected and recently a disturbance of this type was observed at 9500 Mc/s. (3.16 cm.)(12).

A number of investigations have also been made of the brightness distribution across the solar disk, with somewhat conflicting results. Dicke and Beringer(7) found from eclipse measurements that the brightness was approximately uniform at 1.25 cm.; on the other hand, Piddington and Minnett(13), using an interference method, detected a small but significant limb-brightening at this wavelength. Sander(8) deduced that most of the 3.2 cm. radiation came from near the edge of the sun's disk; this result, however, was based on eclipse observations whose accuracy was limited by adverse weather conditions. Finally, Covington(14) obtained eclipse results at 10.7 cm. which indicated a fairly uniform general brightness, but also showed very clearly the presence in the neighbourhood of visible sunspots of small areas brighter than the rest of the disk.

The object of the work described in this paper was to obtain reasonably accurate and extensive data on solar radiation at a wavelength of 3.18 cm., as part of a programme of observations at a number of wavelengths in the microwave spectrum(13, 15). In order to detect any slow variations in intensity, the measurements were extended over a period of three months. The results clearly showed that fluctuations of the type observed at the longer microwave wavelengths also occur to an appreciable extent at 3.18 cm. It was possible to deduce from these observations a base-level of intensity, corresponding to the equivalent temperature of a "quiet" sun. In addition to the precise measurements of intensity, the radiation level was continuously recorded for several hours of each observing day. Several disturbances were found in this way.

Finally, the solar eclipse of November 1, 1948 was used to investigate the distribution of brightness across the sun's disk.\* In Sydney, this eclipse was partial, 0.63 of the sun's diameter being covered at the maximum phase. The results obtained showed that at 3.18 cm. the brightness of the sun's disk is nearly uniform, but increases somewhat from the centre to the limb.

\* Observations were also made at a wavelength of 10 cm. by J. H. Piddington and J. V. Hindman, and at 50 cm. by W. M. Christiansen, D. E. Yabsley, and B. Y. Mills. The results of these measurements are being published separately.

## II. APPARATUS

The equipment is based on principles described by Dicke(16) and except for the radio-frequency components, is identical with that used at 1.25 cm. in connexion with lunar observations(17). The paraboloid reflector is 44 inches in diameter and is mounted on a polar axis about which it may either be turned by hand or driven by a synchronous motor to follow the path of the sun. A telescope is used for accurate alignment of the radio beam on the sun, an important consideration because of the narrowness of the beam ( $2^\circ$  between half power points). To ensure accurate following in the event of cloudy conditions a monitoring device compares the speed of rotation of the aerial with an accurate clock and indicates when the error in alignment exceeds a predetermined value. A horn feed at the focus is connected by waveguide to a box containing the modulating disk rotating at 25 c/s., the balanced crystal mixer, the klystron local oscillator, and the initial stages of the intermediate frequency amplifier. The latter operates at 30 Mc/s. and has a bandwidth of 7.5 Mc/s. between the half power points. The radio-frequency bandwidth is relatively large so that the receiver responds to all radiation in two channels of this width, 30 Mc/s. above and below the local oscillator frequency of 9428 Mc/s.

The sensitivity of the receiver is determined by the usual method of replacing the aerial by a matched termination which can be heated through a measurable temperature interval. The absolute temperature at which the termination emits an amount of radiation equal to that received by the aerial is called the equivalent aerial temperature. The sensitivity of the receiver used in the present experiments is such that a change in aerial temperature of  $1.5^\circ\text{K}$ . produces an output signal equal to the root mean square of the random fluctuations. The possibility of an error due to second harmonic response in the receiving system has been pointed out by Lehaney and Yabsley(11). A correction is necessary if the signals from the aerial and from the calibrating termination do not contain the same proportion of second harmonic power. This effect was investigated by means of an iris, whose transmission for both fundamental and harmonic had been measured. In this way the correction to be applied to measured values of aerial temperature was found to be +4 per cent.

A direct measurement of the power gain of the aerial was made by transmission and comparison tests using two horns(18). The technique involves only the measurement of relative power levels, and the piston attenuator of a 3.2 cm. signal generator was used for this purpose. The power gain determined by this method was 36.2 db. relative to an isotropic radiator.

It is estimated that systematic inaccuracies in the calibration of the apparatus introduce a probable error of  $\pm 6$  per cent.

## III. EXPERIMENTAL METHOD

### (a) *Regular Observations*

Observations of intensity were commenced in October 1948 and were made almost daily, weather permitting, between November 24, 1948 and March 1, 1949. On each observing day two kinds of measurements were carried out.



Firstly, the received intensity was measured as accurately as possible ; secondly, the intensity was recorded continuously though less accurately for several hours in order to look for evidence of disturbed conditions.

For accurate intensity measurements the method was to set the aerial beam ahead of the sun in hour angle and to record the change in receiver output as the sun passed through the beam. This procedure was repeated for a number of declination angles, as set by the sighting telescope, until a maximum deflection was established. The sensitivity of the equipment was determined before and after each series of measurements and the aerial temperature  $T_a$  corresponding to this deflection was calculated. The equivalent black-body temperature  $T_s$  of the sun is related to  $T_a$  by the equation

$$T_s = \frac{4\pi}{G\Omega_s} \cdot T_a,$$

where  $G$  is the maximum power gain of the aerial compared with an isotropic radiator (36.2 db.), and  $\Omega_s$  is the solid angle subtended by the visible sun. Values of  $\Omega_s$  at the time of each observation were computed from the solar diameters given in the *Nautical Almanac*.

The correction for atmospheric absorption is small at 3.18 cm. and could not be measured by the technique previously used at 1.25 cm.(13). It was possible, however, to measure it by observing the variation of received solar intensity over a large range of zenith angles. This procedure was too lengthy to be carried out frequently but since the absorption at 3.18 cm. is due mainly to the oxygen of the atmosphere rather than to water vapour, the value should remain approximately constant. From measurements of this type the absorption along a vertical path was found to be about 1 per cent., increasing to 2 per cent. at a zenith angle of 60°. Since all observations were made at zenith angles lying within this range, a correction of 1.5 per cent. was added to the readings.

The method of obtaining the continuous record of intensity was simply to drive the aerial so that it followed the sun for several hours. At quarter-hour intervals during this period the aerial was turned from the sun to the sky nearby (7.5° from the sun in hour angle) to establish the output level corresponding to zero solar radiation.

### (b) *Eclipse Observations*

The circumstances of the eclipse of November 1, 1948 for an observer in Sydney are shown in Figure 1, all times being Eastern Australian Standard Time (10 hours ahead of Greenwich Mean Time). First optical contact was at 1647 hours and maximum phase at 1745 hours. Since sunset occurred before the end of the eclipse, a complete series of measurements could not be made.

A continuous record of the received intensity was commenced several hours before the eclipse and continued until sunset. The sky was almost completely clear of cloud and it was possible to keep a continuous check on the alignment of the beam by means of the sighting telescope. The sensitivity of the equipment was measured several times before and once after the eclipse. The values obtained immediately before and after the eclipse period are equal within the



accuracy of the measurement (about 1 per cent.). The output level corresponding to zero solar radiation was measured each quarter-hour by turning the beam  $7.5^\circ$  east of the sun in hour angle for one minute. Owing to the low altitude of the sun during the observations, these "background" readings increased rapidly as the horizon was approached. This was due firstly to the increasing amount of radiation accepted by the aerial from the ground and secondly to the increased radiation from the atmosphere along the line of sight. The background readings measured  $7.5^\circ$  east of the sun were therefore less than the true values in the direction of the sun. To allow for this the combined ground and atmospheric radiations were measured before and after November 1

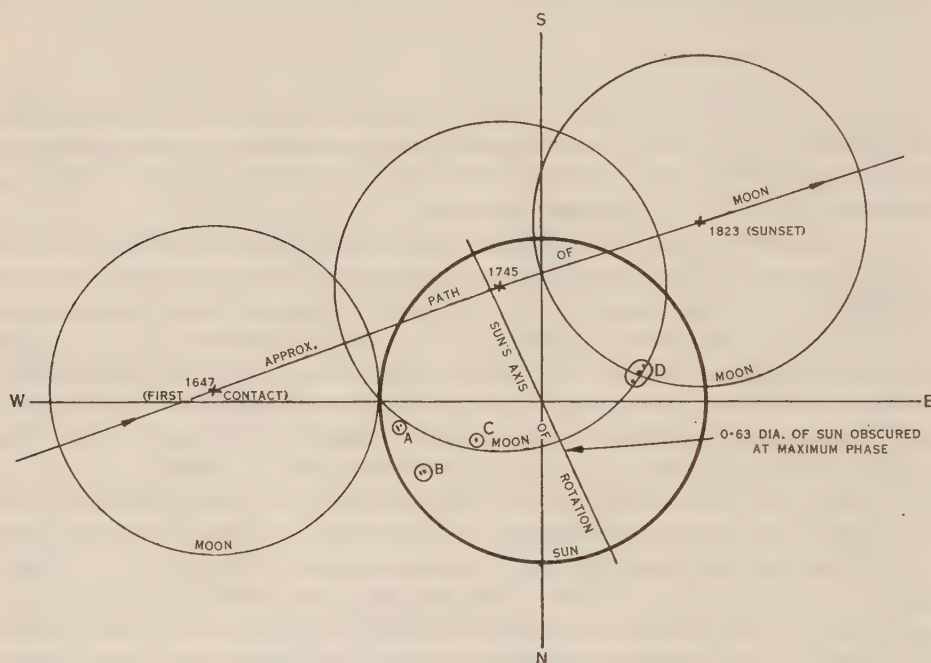


Fig. 1.—Relative positions of the sun and moon during the partial solar eclipse of November 1, 1948 at Sydney (times are Eastern Australian Standard Time).

by moving the aerial beam along the path followed by the eclipsed sun while the sun was in another part of the sky. The resulting curve was then used to correct the background readings obtained during the eclipse. The correction was relatively small before maximum phase but became quite large thereafter. Hence only those observations taken during the first half of the eclipse have been used.

By subtracting the corrected background values from the readings measured with the aerial directed at the eclipsed sun, the received intensity at the aerial was obtained corrected for ground and atmospheric radiation and for zero drifts. A final correction was necessary to allow for atmospheric absorption. This was found directly by measuring the received intensity as the sun traversed the same path on days before and after November 1, 1948.

## IV. RESULTS AND INTERPRETATION

(a) *Radiation from the "Quiet" Sun and Increases Associated with Sunspots*

The results of daily measurements of intensity over a period of three months are plotted in Figure 2 (curve *a*) as the equivalent black-body temperature of the sun. It will be noticed that the temperature varies with time over a total range of approximately 50 per cent. Similar but larger variations have previously been reported at wavelengths of 10 cm. and longer(10, 11) but no measurable changes appear to occur at 1.25 cm.(13).

Since the observations at longer wavelengths show that this variation is associated with the presence of sunspots, the 3.18 cm. results have been examined for correlation with sunspot area and sunspot number. The temperature has been found to correlate much more closely with the area than with the number.

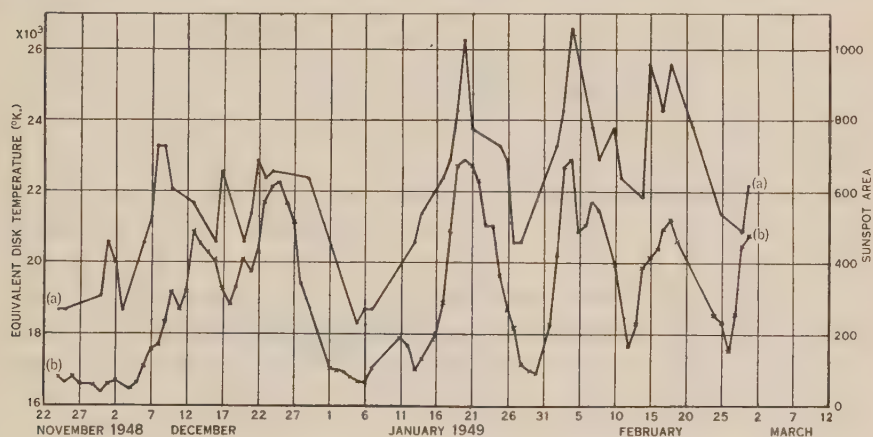


Fig. 2.—Variation with time of equivalent disk temperature and sunspot area.

Curve (a): Equivalent disk temperature.

Curve (b): Sunspot area.

In Figure 2 a curve (b) of sunspot area is shown, the unit of area used being  $10^{-5}$  of the area of the visible disk of the sun. The coefficient of correlation between these variables is 0.79. This marked correlation is further illustrated in Figure 3 in which the equivalent disk temperature is plotted against the sunspot area for the corresponding day. The most probable linear relationship between the two quantities has been calculated by the method of least squares and is represented by the line *AB* in the figure. The background level corresponding to the "quiet" sun is then found to be 19,300 °K. The probable error in this figure is estimated to be  $\pm 7$  per cent., of which  $\pm 1$  per cent. is due to random variations and  $\pm 6$  per cent. is the systematic error in the calibration of the apparatus. This temperature is to be compared with Southworth's figure of 18,000 °K. at 3 cm.(1) and with Sander's value of 22,000 °K. at 3.2 cm.(8). An exact comparison is not possible, however, because the degree to which their results were affected by the sunspot component of the radiation is not known;

the values would probably tend to be higher than the background level. It might also be necessary to take into account possible variation of the level with the progress of the 11-year sunspot cycle.

The slope of the line  $AB$  in Figure 3 is  $8^{\circ}\text{K.}$  per unit sunspot area. If the sources of the increased radio emission have an area equal to that of the visible sunspots, the equivalent temperature of these sources would be  $8 \times 10^5^{\circ}\text{K.}$  However, little is known about the mechanism of sunspot emission and this result is largely speculative. The percentage variation of temperature with sunspot area at  $3.18\text{ cm.}$  is very much less than at the longer microwavelengths.

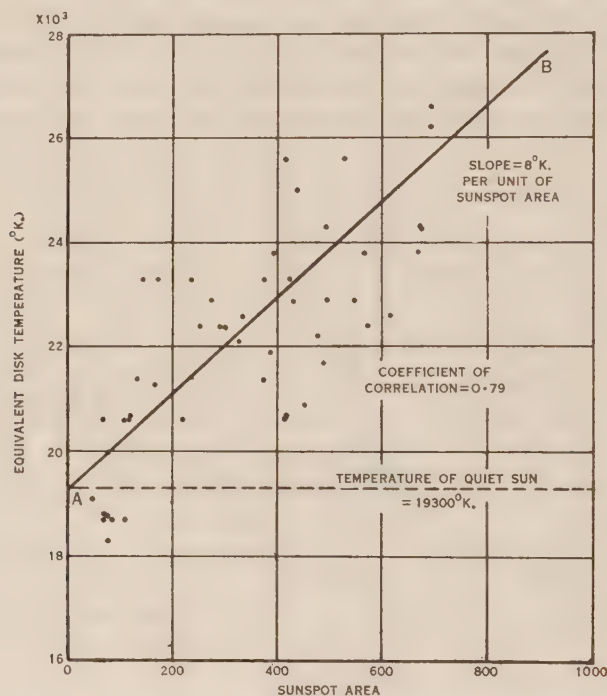


Fig. 3.—Relation between sunspot area and equivalent disk temperature at  $3.18\text{ cm.}$  wavelength between November 1948 and March 1949.

From the figures given above, it is found that the temperature increases by an amount equal to 4 per cent. of the temperature of the “quiet” sun for each hundred units of sunspot area. Data collected by Pawsey and Yabsley(19) indicate that the corresponding figures at  $10.7\text{ cm.}$ ,  $25\text{ cm.}$ , and  $50\text{ cm.}$  are 40, 38, and 24 per cent. respectively.

#### (b) Observations of Bursts

As explained above, continuous recording was carried out on each observing day in order to look for evidence of solar disturbances. It is estimated that changes of received intensity greater than 5 per cent. would be distinguishable from random instrumental fluctuations (caused mainly by variations in the

alignment of the aerial beam). Disturbances of this magnitude at a wavelength of 3.18 cm. are rare and were only observed on five occasions in a total observing time of 230 hours. On each of these occasions the equipment was unattended so that it was not possible to prove directly by turning the aerial away from the sun that the disturbances were of solar origin. Comparison has therefore been made with records of solar radiation obtained at other wavelengths and with other phenomena of solar origin.

The observed increases in intensity ranged from 6 per cent. to over 30 per cent. and lasted for periods of the order of a minute. Details of these disturbances are given in Table 1. The sunspot area was near a maximum in each case, with the exception of November 1 when it was very low. In two instances the bursts were also detected at other centimetre wavelengths—on February 16 at 50 cm. and on February 18 at 10 cm. In both cases the percentage increase in received intensity was greater at 3.18 cm. than at the longer wavelength.

A comparison has also been made with records of observations at 1.5, 3, 3.5, and 5 m. On November 1 and February 16 well-defined bursts were observed at some of these wavelengths at times within a few minutes of the 3.18 cm. disturbances. On January 20 and February 4 the activity at metre wavelengths was so high all day that it masked individual bursts. In contrast to these results, the series of large 3.18 cm. bursts on February 18 was not accompanied by any measurable disturbance at wavelengths greater than 10 cm.

Finally, information on radio fadeouts and solar flares\* was examined. A definite correlation was obtained in one case: the microwave bursts on February 18 were accompanied by a radio fadeout commencing at 1220 E.A.S.T. On January 20 no fadeout occurred but five flares were observed between 1058 and 1550 hours E.A.S.T.; however, the nearest of these to the burst observed on that date finished 21 minutes before the start of the burst. On three other occasions, either flares or fadeouts or both occurred without the appearance of a 3.18 cm. disturbance.

### *(c) Eclipse Observations*

The relative received intensity for a period commencing one and a half hours before visual contact and ending just after maximum phase is shown in Figure 4 (curve *a*). Each point represents the mean intensity over a two-minute interval. Curve (*b*) shows how the intensity from the uneclipsed sun decreased owing to atmospheric absorption over a similar period on days before and after November 1. Both curves have been corrected for the background effect discussed in Subsection III (*b*). It will be observed that the eclipse curve in the period preceding first contact shows fluctuations in level amounting to several per cent. These are instrumental and believed to be caused partly by a strong variable wind which at times made accurate alignment of the narrow beam difficult.

\* These records were supplied by the Commonwealth Observatory, Canberra. The flare data are incomplete as continuous observations are not made.



TABLE 1  
BURSTS OBSERVED AT 3.18 CM. WAVELENGTH AND CORRELATION WITH RECORDS AT LONGER WAVELENGTHS

Date	3.2 cm. data					Bursts Recorded at Other Wavelengths						
	Time of Com- mencement (hr. E.A.S.T.)	$T_1$ ( $^{\circ}\text{K.} \times 10^{-3}$ )	$T_2$ ( $^{\circ}\text{K.} \times 10^{-3}$ )	Increase (%)	Duration (min.)	10 cm.	25 cm.	50 cm.	1.5 m. (200 Mc/s.)	3 m. (100 Mc/s.)	3.5 m. (85 Mc/s.)	5 m. (60 Mc/s.)
1948												
Nov. 1 1949	1308½	18.9	20.5	9	2	—	—	—	1311-1312 $I=14$	—	1317 $I=60$	1310 $I=7$
Jan. 20	1131	26.2	28.7	12	½	—	—	—	$a$	$a$	—	—
Feb. 4	1453½	26.6	> 33.5	> 26	2	—	—	—	$a$	$a$	—	—
Feb. 16	1321¾	24.9	> 33.2	> 33	3	—	—	1320-23 6%	1321-1326 $I=3.6$	1318-1323 $I=350$	—	1317-1330 $I=60$
Feb. 18	1219		30.8	20	Total duration 9 min.	1219, 12%	$b$	$b$	$b$	$b$	—	$b$
	1222	25.6	> 31.2	> 22		1222, 16%						
	1227		> 31.2	> 22		1227, 10%						

*Explanatory notes:* At wavelengths up to 50 cm. the maximum percentage increase in intensity is given. At longer wavelengths the quantity tabulated is the maximum smoothed power flux density,  $I$ , during the disturbance (unit is  $10^{-2}$  watts metre $^{-2}$  (c/s.) $^{-1}$ ). A blank in the table indicates that no equipment of the specified wavelength was operating at the time.

$T_1$  = equivalent black-body temperature before and after disturbance,

$T_2$  = highest equivalent black-body temperature during burst (where recorder ran off-scale, temperature corresponding to full-scale deflection is given),

$a$  = high activity throughout day,

$b$  = no disturbance near time of 3.18 cm. bursts.

On the day of the eclipse there were six sunspot groups, the total area having the low value of 85. The eclipsed groups were located as shown in Figure 1. These are marked *A*, *B*, *C*, and *D* and their times of covering and uncovering are indicated in Figure 4. Any changes due to the covering and uncovering of the spots are too small to be distinguished from instrumental variations. Now the figures given in Subsection IV (*a*) for the increase in radiation caused by unit sunspot area show that about 4 per cent. of the total 3.18 cm. radiation on November 1 would be expected to be due to sunspots. Thus the covering or uncovering of one of the larger groups would change the received intensity by about 1 per cent. and variations of this order would be obscured by the observed instrumental variations.

A downward trend in the intensity curve began seven minutes before first optical contact at which time the value had decreased to 96 per cent. This change may have been instrumental but it is of interest to examine the possibility of its being a real effect in view of the 50 cm. observations made by Christiansen,

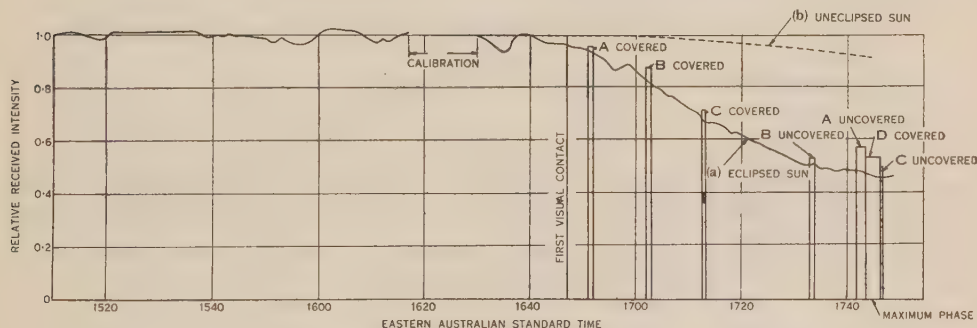


Fig. 4.—Relative received intensity of 3.18 cm. solar radiation.

(a) On November 1, 1948 before and during eclipse.

(b) Mean values for October 26 and November 2, 1948.

Yabsley, and Mills(20) during the same eclipse. By using three equipments they were able to eliminate instrumental changes to a large extent and their results also show a decrease of intensity beginning about seven minutes prior to optical contact. If this effect is to be explained by the eclipsing of radiating levels in the solar atmosphere, such levels would have to be at least 140,000 km. above the sun's visible surface. On theoretical grounds(21) it seems unlikely that any substantial amount of 3.18 cm. radiation would originate from such a high level in the "quiet" sun. However, the result could possibly have been caused by a localized emitting region extending from the sun's limb.

To examine the general distribution of brightness over the solar disk, the experimental results have been compared with those to be expected from various distributions having circular symmetry. The theoretical variation of received intensity is shown in Figure 5 for the following brightness distributions :

(a) Uniform over the visible disk.

(b) Concentrated entirely at the circumference of the visible disk.

- (c) Uniform over a disk with a radius 1.1 times that of the visible disk.
- (d) Having 74 per cent. of the total received energy originating in a source of type (a) and 26 per cent. in a source of type (b).

The experimental points shown in Figure 5 are derived from the experimental curve (a) of Figure 4 by smoothing out the short period fluctuations and applying the correction for atmospheric absorption from curve (b). The points have also been adjusted to a value of unity at the instant of optical contact. This is correct if the initial decrease of intensity was due to the covering of a localized source of energy on the solar limb. The same procedure also seems justified if the effect is considered to be the result of an instrumental change because there is insufficient data on which to base a more elaborate correction.

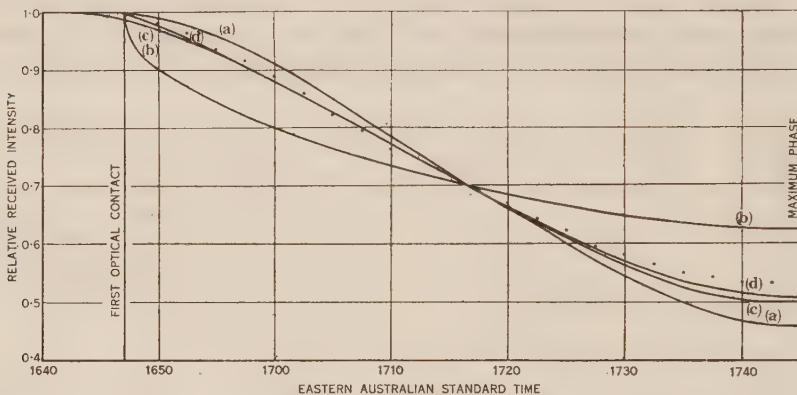


Fig. 5.—Comparison of measured variation of relative received intensity during eclipse with theoretical results for various distributions of brightness over the sun's disk.

.... Smoothed experimented points.

Brightness distribution :

- (a) Uniform over visible disk.
- (b) Concentrated entirely at circumference of visible disk.
- (c) Uniform over disk with radius 1.1 times that of visible disk.
- (d) Seventy-four per cent. of total received energy originating in source of type (a), 26 per cent. in source of type (b).

It is evident from Figure 5 that the experimental results differ significantly from those to be expected from a uniform disk of visual size. They are in reasonable agreement with brightness distributions of either type (c) or type (d) for which the curves are very similar except near first contact and at maximum phase. The uncertainty in the experimental results is sufficient to prevent distinctions being drawn between the two brightness distributions. The measurements therefore are not conclusive regarding the limb-brightening effect predicted by theory(5, 6) although they are consistent with the existence of such an effect.

## V. CONCLUSIONS

The results of regular observations over a period of three months between November 24, 1948 and March 1, 1949 indicate a value of 19,300 °K. for the equivalent black-body temperature of the "quiet" sun at 3.18 cm. The probable error in this temperature is  $\pm 7$  per cent. The temperature increased by about 8 °K. for each unit increase in sunspot area.

Bursts of radiation at this wavelength are rare. In a total observing time of about 230 hours, evidence of bursts was obtained on five occasions. These disturbances lasted for periods of the order of a minute and corresponded to increases in the equivalent disk temperature of from 6 per cent. to over 30 per cent.

Observations made during the eclipse of November 1, 1948 showed that the temperature distribution over the disk was consistent with either of the following brightness distributions:

- (i) Combination of a uniform disk of optical size emitting 74 per cent. of the total energy together with a thin bright ring around the circumference.
- (ii) Uniform over a disk 1.1 times the optical size.

## VI. ACKNOWLEDGMENTS

The authors wish to acknowledge the advice of Dr. J. H. Piddington in planning the observations, and the assistance of Mr. J. V. Hindman in developing the equipment. Data on sunspots, flares, radio fadeouts, and 200 Mc/s. observations were made available through the courtesy of the staff of the Commonwealth Observatory, Mt. Stromlo, Canberra. We are also indebted to Miss R. Payne-Scott for the records of 60, 85, and 100 Mc/s. observations.

## VII. REFERENCES

- (1) SOUTHWORTH, G. C.—*J. Franklin Inst.* **239**: 285-97 (1945); and correction **241**: March 1946.
- (2) APPLETON, E. V., and HEY, J. S.—*Phil. Mag.* **37**: 73-84 (1946).
- (3) MCCREADY, L. L., PAWSEY, J. L., and PAYNE-SCOTT, R.—*Proc. Roy. Soc. A* **190**: 357-75 (1947).
- (4) GINSBURG, V. L.—*C.R. Acad. Sci. U.R.S.S.* **52**: 487-90 (1946).
- (5) MARTYN, D. F.—*Nature* **158**: 632-3 (1946).
- (6) MARTYN, D. F.—*Proc. Roy. Soc. A* **193**: 44-59 (1948).
- (7) DICKE, R. H., and BERINGER, R.—*Astrophys. J.* **103**: 375-6 (1946).
- (8) SANDER, K. F.—*Nature* **159**: 506 (1947).
- (9) REBER, G.—*Nature* **158**: 945 (1946).
- (10) COVINGTON, A. E.—*Proc. Instn. Radio Engrs.* **36**: 454-7 (1948).
- (11) LEHANY, F. J., and YABSLEY, D. E.—*Aust. J. Sci. Res. A* **2**: 48-62 (1949).
- (12) SCHULKIN, M., HADDOCK, F. T., DECKER, K. M., MAYAR, C. H., and HAGEN, J. P.—*Phys. Rev.* **74**: 840 (1948).
- (13) PIDDINGTON, J. H., and MINNETT, H. C.—*Aust. J. Sci. Res. A* **2**: 541 (1949).
- (14) COVINGTON, A. E.—*Nature* **159**: 405-6 (1947).
- (15) PIDDINGTON, J. H., and HINDMAN, J. V.—*Aust. J. Sci. Res. A* **2**: 526 (1949).
- (16) DICKE, R. H.—*Rev. Sci. Instrum.* **17**: 268-75 (1946).
- (17) PIDDINGTON, J. H., and MINNETT, H. C.—*Aust. J. Sci. Res. A* **2**: 63-77 (1949).
- (18) CUTLER, A., KING, A. P., and KOCH, W. E.—*Proc. Instn. Radio Engrs.* **35**: 1462-71 (1947).
- (19) PAWSEY, J. L., and YABSLEY, D. E.—*Aust. J. Sci. Res. A* **2**: 198-213 (1949).
- (20) CHRISTIANSEN, W. N., YABSLEY, D. E., and MILLS, B. Y.—*Aust. J. Sci. Res. A* **2**: 508 (1949).
- (21) SMERD, S. F.—*Aust. J. Sci. Res. A* **3**: 34-59 (1950).



# DEFORMATION AND RECRYSTALLIZATION OF ALLOYS CONTAINING TWO PHASES

By L. M. CLAREBROUGH\*

[Manuscript received November 3, 1949]

## Summary

Using mainly silver-magnesium alloys, the deformation of the phases was determined from observations of slip lines and recrystallization behaviour. The relative deformation of the two phases is influenced both by the total deformation of the alloy and by the volume fractions of the phases. When very little of the hard  $\beta$  phase is present most of the deformation occurs in the soft  $\alpha$  crystals and there is additional flow of the matrix around the hard particles. If the proportion of the  $\beta$  phase is not greater than 30 per cent. by volume the soft  $\alpha$  deforms more than the  $\beta$  for reductions of up to 60 per cent., but for heavier reductions the two phases deform more uniformly. When both phases are present in about equal proportions they deform to the same extent.

The recrystallization behaviour of the phases in a duplex alloy is affected by the state of order of the phases, and by precipitation occurring in one of the phases on annealing the deformed alloy. Experiments were designed to show the influence of various factors.

An alloy that is ordered before deformation recrystallizes at a lower temperature than an alloy disordered before deformation. It is shown that this behaviour is not due to a difference in work hardening characteristics.

On annealing silver-magnesium and copper-zinc alloys quenched before deformation, precipitation of  $\alpha$  occurs in the  $\beta$  crystals and the recrystallization temperature of the  $\beta$  is higher in these alloys than in slowly cooled alloys. It is shown that precipitation relieves internal stresses in the deformed matrix and thus causes the increase in recrystallization temperature.

The general case of recrystallization of the phases in slowly cooled and quenched silver-magnesium alloys is interpreted in terms of these results.

The inhomogeneity of deformation of the phases is indicated by their mode of recrystallization.

## I. INTRODUCTION

In a recent paper by Honeycombe and Boas(1) the deformation and recrystallization of duplex brass were described. Their results were complicated because the  $\beta$  phase becomes disordered at 454 °C. and the rate of cooling through the transformation range before deformation affects the recrystallization characteristics of this phase.

In a search for an alloy system without the complication of an order-disorder transformation and with a composition range in which two ductile phases existed at room temperature, the silver rich silver-magnesium alloys were thought to be suitable. In this system the  $\beta$  phase is ordered at all temperatures up to the melting point. The existence of a superlattice in the  $\alpha$  phase at 25 atomic

\* Division of Tribophysics, C.S.I.R.O., University of Melbourne.

per cent. magnesium was suggested by Andrews and Hume-Rothery(2), but this was not detected in a subsequent investigation by Letner and Sidhu(3). During the present investigation the superlattice was found and details of the structure and transformation will be published elsewhere.

A portion of the silver-magnesium phase diagram is shown in Figure 1. The boundaries of the  $\alpha + \beta$  phase field are not vertical, and there is an increase in the solubility of magnesium in the  $\alpha$  phase and a decrease in the solubility of magnesium in the  $\beta$  phase as the temperature increases, with consequent change in the proportion of the phases.

Most of the experimental work was done with duplex silver-magnesium alloys, but certain points were clarified by experiments with copper-zinc alloys and a copper-iron alloy.

## II. EXPERIMENTAL

All the silver-magnesium alloys were prepared from commercial magnesium and fine silver and analysed for magnesium and silver. The compositions of the materials used are given in Table 1. Except in one case, the total impurity content never exceeded 0.2 per cent. The silver and magnesium were melted together in a salamander crucible under a commercial magnesium flux, the main

TABLE 1  
COMPOSITION OF MATERIALS USED IN THE INVESTIGATION

	Silver (%)	Magnesium (%)	$\alpha$ - $\beta$ Brass (%)	$\beta$ Brass (%)	Copper-Iron Alloy (%)
Silver .. ..	99.9*				
Magnesium .. ..	Faint trace	99.88*			
Copper .. ..	Trace	<0.01	58.5	50.6	79.82
Zinc .. ..	Faint trace	Trace	41.3*	49.4*	
Lead .. ..	<0.1	<0.01	0.07	0.01	
Nickel .. ..			Trace		Trace
Iron .. ..	Faint trace	<0.05	0.13	Trace	20.0
Manganese .. ..		<0.05			Trace

\* By difference.

constituent of which was anhydrous magnesium chloride, and the melt was cast into  $\frac{1}{2}$ -in. diameter chill moulds. To remove any porosity the castings were remelted in  $\frac{1}{2}$ -in. diameter graphite crucibles in which they were allowed to solidify. The resulting ingots were turned to  $\frac{3}{8}$ -in. diameter, swaged to  $\frac{3}{16}$ -in. diameter with one intermediate anneal, annealed at 600 °C. for six days, and then slowly cooled to room temperature at a rate of 30 °C. per hour. For all homogenizing treatments the alloys were sealed in evacuated pyrex glass tubes.

Wire drawing was the main method of deformation, but rolling and swaging were also used in some cases. For X-ray and microscopic studies of recrystallization two identical specimens of wire were annealed together in a thin pyrex glass tube immersed in a salt bath. A standard annealing time of 30 minutes was used throughout this work. The specimens for X-ray examination were etched down to remove the highly deformed outer layer and back reflection diagrams were made using a 1-mm. pinhole and copper radiation.

The microscopic examination was done on longitudinal sections of the drawn wires, and for every annealing temperature microscopic and X-ray examinations were combined to determine the state of recrystallization. Electrolytic polishing of these alloys was not successful and all specimens were polished mechanically. Different reagents were used to etch the phases and it was not found possible to develop the detailed structures of the phases simultaneously, as reagents that effectively etched the  $\alpha$  phase destroyed all detail in the  $\beta$  phase. The following etchants were used:

$\alpha$ Phase	$\beta$ Phase
5 ml. nitric acid, conc. 25 ml. hydrochloric acid, conc. 50 ml. water	25 g. chromium trioxide 0.2 g. sodium sulphate 100 ml. water

Microscopic examination was more reliable than X-ray examination as in many cases, owing to fine recrystallized grain size, continuous Debye-Scherrer circles were obtained from a phase that was fully recrystallized. The combination of these two methods of examination enabled the state of recrystallization of the phases to be classified as just started, started, partly, nearly fully, and fully.

The metallographic technique for the commercial copper-zinc alloys used was the same as that described by Honeycombe and Boas(1).

The copper-iron alloy, used to show certain features of the deformation of duplex alloys, was prepared by melting electrolytic copper and "Armco" iron together in an alundum crucible using a high-frequency induction furnace. The alloy was prepared under reduced pressure (approximately 0.1 mm. Hg) and allowed to solidify in the crucible. Sufficient oxygen was present to give rise to discrete particles of iron oxide in the cast structure.

### III. RESULTS

#### (a) *Deformation of Silver-Magnesium and Copper-Iron Alloys*

The deformation of the phases in the silver-magnesium alloys was followed by microscopic examination of the development of slip bands on polished and etched faces of small compression specimens.

In an alloy, slowly cooled from 600 °C., containing approximately 25 volume per cent. of the hard  $\beta$  phase, slip commenced in the  $\alpha$  phase after slight



compression (approximately 1 per cent.). After 10 per cent. compression the  $\alpha$  phase showed very heavy slip, but there was practically no sign of deformation markings in the  $\beta$  phase (Plate 1, Fig. 1). In some instances, however, slip lines from the  $\alpha$  crystals continued for a short distance into the  $\beta$  crystals without change in direction (Plate 1, Fig. 2). Since only a small fraction of the  $\alpha$  is formed from the  $\beta$  on cooling the alloy from 600 °C., it seems unlikely that the continuation of slip from the  $\alpha$  into the  $\beta$  can be attributed to a crystallographic relationship between the two phases. It may be possible, however, for slip lines in the  $\alpha$  phase running up to an  $\alpha$ - $\beta$  boundary to initiate slip in the  $\beta$  crystals across this boundary. In these regions the  $\beta$  crystals are obviously more heavily deformed than in other parts.

In order to observe the influence of more heavy working on the relative deformation of the phases, specimens were cut from the alloy after 60 per cent. reduction in thickness by rolling, and then polished and etched. After 60 per cent. reduction, followed by an additional slight compression, slip still commenced in the  $\alpha$  phase, but after 10 per cent. compression both phases showed slip and surface rumpling although the  $\alpha$  phase appeared more heavily deformed than the  $\beta$  (Plate 1, Fig. 3). In the  $\alpha$  phase the slip bands were of a wavy nature.

Similar behaviour was observed in the alloy after 75 per cent. reduction by rolling followed by 10 per cent. compression. Both phases appeared to be equally deformed and pronounced slip bands were visible in the  $\beta$  phase crystals (Plate 1, Fig. 4).

These results indicate that in an alloy containing approximately 25 volume per cent. of the  $\beta$  phase, the  $\alpha$  phase is more heavily deformed than the  $\beta$  in the initial stages of the deformation, but as the degree of deformation increases the  $\beta$  phase takes part to a greater extent.

It is to be expected that the relative extent of deformation of the two phases in a duplex alloy is influenced by the proportion of the phases. To show this influence, two silver-magnesium alloy specimens containing 25 and 70 per cent. by volume of the  $\beta$  phase respectively, were given 5 per cent. compression. In the first alloy the  $\alpha$  phase showed heavy slip but there was no sign of deformation in the  $\beta$  (Plate 1, Fig. 5). However, in the alloy containing 70 per cent. of the  $\beta$  the appearance of the slip bands showed both phases to be deformed to about the same extent (Plate 1, Fig. 6). Thus the degree of deformation of the  $\beta$  phase increases as its proportion is increased.

In these alloys slip line observations did not indicate any inhomogeneity of deformation in the  $\alpha$  phase near the  $\beta$  crystals. A marked inhomogeneity may be expected if a small number of hard particles is distributed in a soft matrix. This effect can be demonstrated in a copper-iron alloy containing dispersed particles of iron oxide. The structures of the alloy as cast and after 90 per cent. reduction in thickness by rolling are shown in Plate 2, Figs. 7, 8, and 9. During deformation the small round grains composing the iron dendrites form long streamers. These show that flow in the matrix is disturbed around the hard, brittle iron oxide particles and that there is additional compression in these regions. The type of deformation illustrated by this alloy is that



predicted by Unkel(4). Thus when the proportion of the hard second phase is small the deformation occurring in the matrix around the hard particles is greater than in other parts of the specimen.

The relative deformation of the hard and soft phases in a duplex alloy is influenced both by the total deformation and by the proportion of the phases present. Certain features of the deformation will become clearer when considering the mode of recrystallization of the phases, as it can be assumed that recrystallization starts in the most heavily deformed regions of a phase (Section III (d)).

*(b) Influence of Various Factors on the Recrystallization Behaviour*

From the results obtained with brass(1), the recrystallization behaviour of duplex alloys of this type can be expected to be different depending on whether the alloys are quenched from a high temperature or slowly cooled to room temperature before deformation. This difference in behaviour is due to the following factors which are controlled by the heat treatment :

- (i) State of order of the phases.
- (ii) Proportion of the phases.
- (iii) Degree of departure from equilibrium at the recrystallization temperature.
- (iv) Composition of the phases.

Experiments were designed to isolate the first three factors in turn so that their contribution to the general case could be determined.

(i) *State of Order of the Phases.*—In the silver-magnesium alloys the disordered state of the  $\alpha$  phase can be retained by quenching from above 400 °C., but the  $\beta$  phase is ordered at all temperatures up to the melting point. In brass the  $\alpha$  phase is disordered at all temperatures but the  $\beta$  phase is ordered only below 454 °C. Although the disordered state cannot be retained by quenching, the degree of order is said to be less after such a treatment(5).

It has been observed in duplex brass that the recrystallization behaviour of the  $\beta$  phase depends on whether the alloy has been quenched or slowly cooled from above the critical temperature before deformation. The  $\alpha$  phase is not influenced by this treatment. The different work hardening characteristics of the  $\beta$  phase in the two conditions were considered responsible for this behaviour.

In order to obtain a clearer indication of the influence of the degree of order of  $\beta$  brass on its recrystallization behaviour two specimens of single phase  $\beta$  brass strip were annealed for 16 hours at 600 °C. One specimen was quenched, the other slowly cooled to room temperature, and both specimens were rolled to 20 per cent. reduction in thickness. The hardness values are given in Table 2.

The slowly cooled alloy was partly recrystallized after annealing at 400 °C. and almost fully recrystallized at 450 °C., whereas the quenched alloy just started to recrystallize at 450 °C. (Plate 2, Figs. 10 and 11).

To study this effect in the silver-magnesium alloy  $\text{Ag}_3\text{Mg}$  an alloy rod containing 24.7 atomic per cent. magnesium was quenched from 600 °C. so as to produce a disordered structure, and another rod of the same composition

was slowly cooled to room temperature at a rate of 30 °C. per hour so as to produce the ordered structure. The recrystallization behaviour of these two alloys after various reductions by wire drawing is summarized in Figure 2.

TABLE 2\*  
INFLUENCE OF DEFORMATION ON THE HARDNESS OF QUENCHED AND SLOWLY COOLED  $\beta$  BRASS

% Reduction by Rolling	Vickers Hardness Number (10-kg. Load)	
	Quenched	Slowly Cooled
0	147	95
20	217	195

\* Each value is an average of at least three readings.

For deformations up to 50 per cent. reduction in area the start of recrystallization is not influenced greatly by the degree of order of the alloy before deformation, but after 70 and 80 per cent. reduction in area the alloy that was ordered before deformation starts to recrystallize at a temperature 25 °C. lower than the alloy disordered before deformation. The initial hardness values and the work hardening behaviour of these alloys after various reductions by rolling are shown in Table 3.

TABLE 3\*  
INFLUENCE OF DEFORMATION ON THE HARDNESS OF QUENCHED AND SLOWLY COOLED  $\text{Ag}_3\text{Mg}$

% Reduction by Rolling	Vickers Hardness Number (10-kg. Load)	
	Quenched	Slowly Cooled
0	66	100
20	103	129
50	141	159
70	161	172

\* Each value is an average of at least three readings.

The disordered state of the  $\text{Ag}_3\text{Mg}$  superlattice is softer than the ordered state and work hardens more, whereas quenched  $\beta$  brass is harder than slowly

cooled  $\beta$  brass and work hardens less. In spite of this the recrystallization behaviour of the two alloys is the same. Thus the lower recrystallization temperature of the slowly cooled alloys cannot be caused by different work hardening characteristics but must be due to some other factor resulting from the state of order of the alloys.

(ii) *Proportion of the Phases.*—To investigate the influence of the proportion of the phases on their recrystallization behaviour, three silver-magnesium

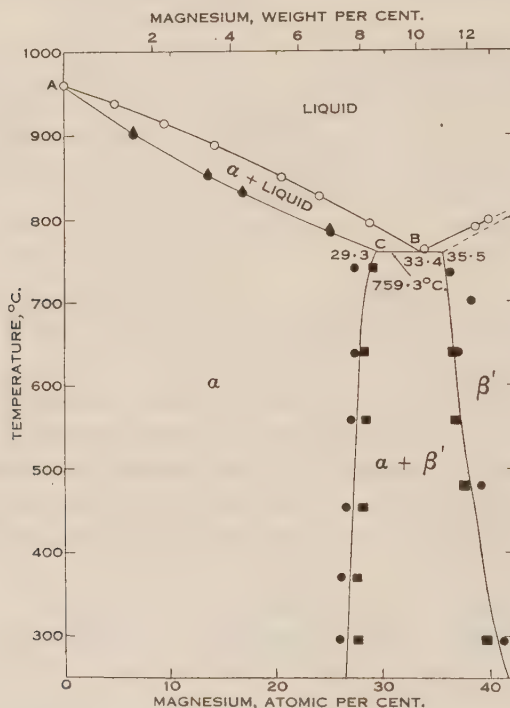


Fig. 1.—Portion of the silver-magnesium phase diagram according to Andrews and Hume-Rothery(2).

alloys containing 10, 40, and 70 per cent. by volume of the  $\beta$  phase were deformed to 22 per cent. reduction in area by swaging.

These alloys had been annealed at 600 °C. and slowly cooled to room temperature before deformation so that in all cases the phases were of the same composition and no precipitation of  $\alpha$  occurred in the deformed  $\beta$  on annealing for recrystallization. Swaging had to be used to deform the alloys, as that containing 70 per cent. of  $\beta$  was too brittle to be drawn. The temperature at which recrystallization started in the  $\beta$  phase was the same for the alloys containing 40 and 70 per cent. by volume of the  $\beta$  phase, but was 25 °C. higher in the alloy containing only 10 per cent. of this phase. The temperature at which recrystallization started in the  $\alpha$  phase appeared to be independent of the proportion of the  $\beta$  phase, recrystallization having started in all cases at 375 °C.

These results are in agreement with slip line observations and indicate that the hard  $\beta$  phase is more deformed as its proportion increases.

(iii) *Degree of Departure from Equilibrium at the Recrystallization Temperature.*—Owing to the change in the phase equilibrium with temperature in the silver-magnesium system, neither slowly cooled nor quenched alloys are in an equilibrium condition at the annealing temperatures used for recrystallization (Fig. 1). On annealing slowly cooled alloys, an increase in the proportion of the  $\beta$  phase occurs by growth of the  $\beta$  crystals, and there is very little precipitation of  $\beta$  in the  $\alpha$  phase. On annealing quenched alloys at any temperature below that of quenching, a decrease in the proportion of the  $\beta$  phase occurs by precipitation of  $\alpha$  in the  $\beta$  crystals.

To study the influence of precipitation in the  $\beta$  phase of alloys quenched before deformation, two alloys, prepared so that the proportion of the  $\beta$  phase was the same in both, were given the same deformation and the same annealing treatment. Their heat treatment before deformation was such that precipitation of  $\alpha$  in  $\beta$  occurred in one alloy on annealing but not in the other. The treatment of the alloys and their recrystallization behaviour are set out in Table 4. The results are shown by the X-ray diagrams and microstructures in Plate 3, Figures 12, 13, and 14. The Debye-Scherrer circle for the  $\beta$  phase of the quenched alloy is quite sharp after annealing at 350 °C., indicating that the internal stresses that had been present in this phase owing to plastic deformation and had given rise to line broadening, were removed.

When specimens of the deformed alloys were annealed at 300 °C. the  $\beta$  phase did not recrystallize in either case and the lines in the X-ray patterns were broad because of the deformation. These lines were less broad, however, in the diagram of the quenched alloy. Precipitation had occurred in the quenched alloy but not in the slowly cooled alloy.

The same changes in phase proportions occur in copper-zinc alloys where the changes in proportion involved are even greater. Here, however, the increase in proportion of the  $\beta$  phase which occurs on heating slowly cooled alloys takes place by precipitation of  $\beta$  in  $\alpha$  rather than by growth of the  $\beta$  crystals. As shown in Table 4, precipitation of  $\alpha$  in the  $\beta$  phase of quenched brass during annealing for recrystallization produces the same result as for the silver-magnesium alloys. The microstructures and X-ray diagrams of these alloys after annealing at 300 °C. are shown in Plate 3, Figures 15, 16, and 17. As with the silver-magnesium alloys, the Debye-Scherrer circle for the  $\beta$  phase of the quenched alloy is quite sharp. Since only one composition of brass was used, the quenched alloy contained a higher proportion of  $\beta$  than the slowly cooled alloy. It has been shown in Section III (b) (ii) that the recrystallization temperature of a hard second phase is lowered as its proportion increases. The fact that the recrystallization temperature of the  $\beta$  phase increases despite an increase in the proportion of the  $\beta$  indicates again that this increase is due to the relief of internal stresses by precipitation. It should be realized, however, that in brass there is also another factor acting in the same direction, i.e. the degree of order of the  $\beta$  phase.



TABLE 4  
INFLUENCE OF PRECIPITATION ON RECRYSTALLIZATION BEHAVIOUR OF SILVER-MAGNESIUM AND COPPER-ZINC ALLOYS

	Silver-Magnesium		Copper-Zinc	
	Alloy 1	Alloy 2	Alloy 1	Alloy 2
Composition (wt. %)	9.7% Mg	9.5% Mg	41.3% Zn	41.3% Zn
Heat treatment	Slowly cooled to 300 °C., held at 300 °C. for 1 hr., and quenched.	Quenched from 600 °C.	Slowly cooled to 300 °C., held at 300 °C. for 4 days, and quenched.	Quenched from 600 °C.
Proportion of phases	Approx. 50%.	Approx. 50%.	Approx. 40%.	Approx. 60%.
Deformation (wiredrawing)	80% reduction in area	80% reduction in area	80% reduction in area	80% reduction in area
Annealing treatment	30 min. at 350 °C.	30 min. at 350 °C.	30 min. at 300 °C.	30 min. at 300 °C.
Recrystallization of $\alpha$ phase	Fully recrystallized	Fully recrystallized	Partly recrystallized	Recrystallization started
Recrystallization of $\beta$ phase	Almost fully recrystallized (Plate 3, Figs. 12 (a) and 13)	No recrystallization (Plate 3, Figs. 12 (b) and 14)	Fully recrystallized (Plate 3, Figs. 15 (a) and 16)	No recrystallization (Plate 3, Figs. 15 (b) and 17)
Precipitation	No precipitation of $\alpha$ in $\beta$ (Plate 3, Fig. 13)	Heavy precipitation of $\alpha$ in $\beta$ (Plate 3, Fig. 14)	No precipitation of $\alpha$ in $\beta$ (Plate 3, Fig. 16)	Heavy precipitation of $\alpha$ in $\beta$ (Plate 3, Fig. 17)

The conclusion is drawn from these experiments that precipitation occurring in the deformed  $\beta$  phases on annealing quenched duplex silver-magnesium and copper-zinc alloys is responsible for the relief of internal stresses in these phases, thus reducing the tendency for recrystallization(6).

(iv) *Composition of the Phases.*—No experiments were carried out to isolate the influence of the composition of the phases on their recrystallization behaviour. In the silver-magnesium alloys it is expected that the recrystallization temperatures of both the  $\alpha$  and  $\beta$  phases increase with increasing magnesium content. The influence of composition, however, is probably small compared with that of the other factors already discussed.

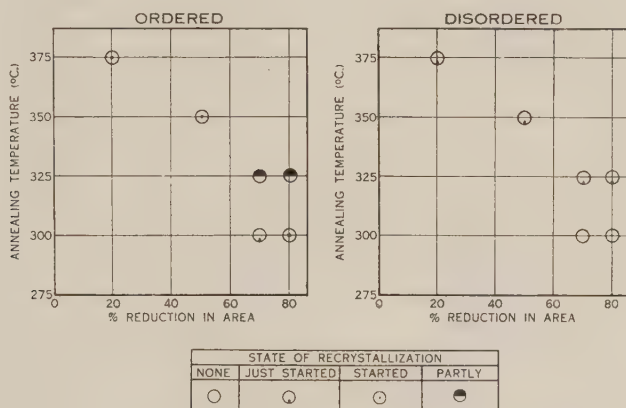


Fig. 2.—Influence of deformation on recrystallization of ordered and disordered  $\text{Ag}_3\text{Mg}$ .

(c) *General Case of Recrystallization in Slowly Cooled and Quenched Duplex Silver-Magnesium Alloys*

To investigate the general case where all the factors discussed above are combined, two sets of silver-magnesium alloy rods were annealed for three days at  $600^\circ\text{C}.$ ; one set was quenched and the other slowly cooled to room temperature at a rate of  $30^\circ\text{C}.$  per hour. The composition of the quenched rods was  $32.7 (\pm 0.1)$  atomic per cent. magnesium and that of the slowly cooled rods  $31.3 (\pm 0.5)$  atomic per cent. magnesium. Both sets of rods were deformed by wire drawing to various extents and were annealed at various temperatures. Their recrystallization behaviour is shown in Figure 3 and typical microstructures and diffraction patterns are given in Plates 4 and 5, Figures 18–26. The main points illustrated by Figure 3 may be summarized as follows:

- (i) Temperature at which recrystallization starts in the phases of duplex silver-magnesium alloys

(1)  $\alpha$  phase

After low deformations: at almost the same temperature in slowly cooled and quenched alloys.

After high deformations: at least  $25^\circ\text{C}.$  lower in the slowly cooled alloys than in the quenched alloys.

(2)  $\beta$  phase

At lower temperatures in slowly cooled alloys than in quenched alloys.

Increasing the amount of deformation from 47 to 90 per cent. lowers the temperature at which recrystallization starts by at least 25 °C. in slowly cooled alloys, whereas in quenched alloys this temperature (400 °C.) is almost independent of deformation.

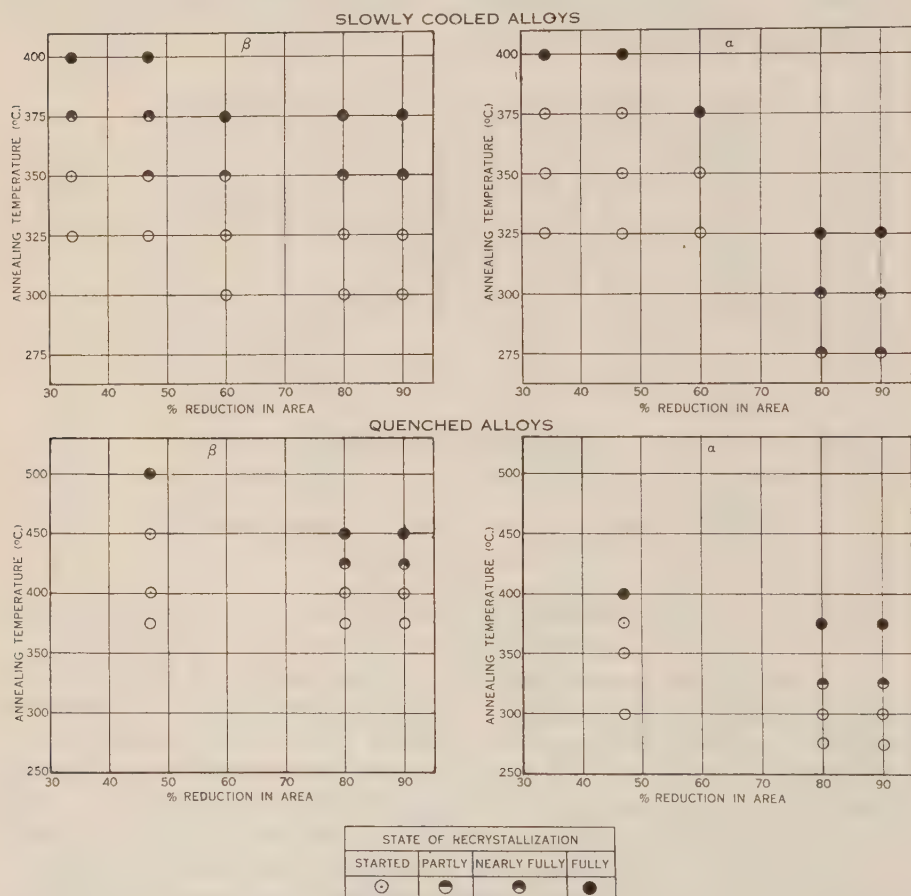


Fig. 3.—Influence of deformation on recrystallization of the phases in slowly cooled and quenched silver-magnesium alloys.

## (ii) Temperature at which recrystallization is complete

(1)  $\alpha$  phase

At lower temperatures in slowly cooled than in quenched alloys.

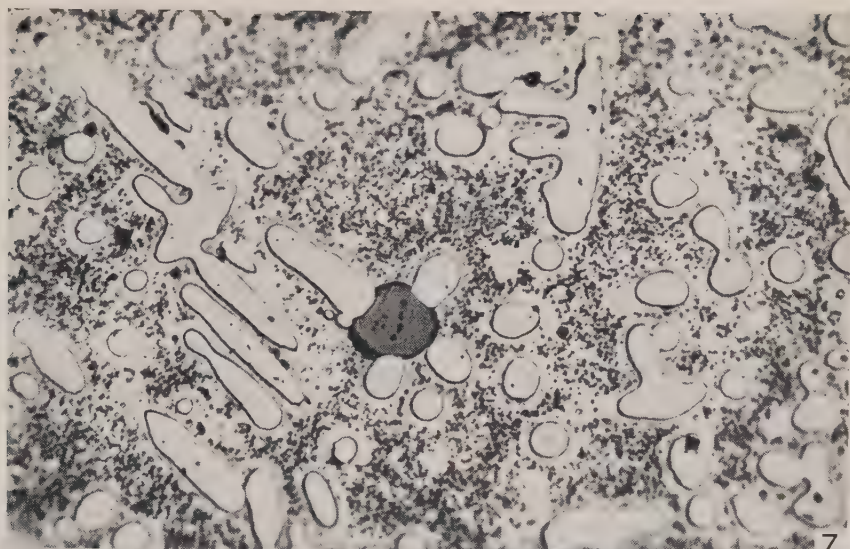
(2)  $\beta$  phase

At much lower temperatures in slowly cooled than in quenched alloys.

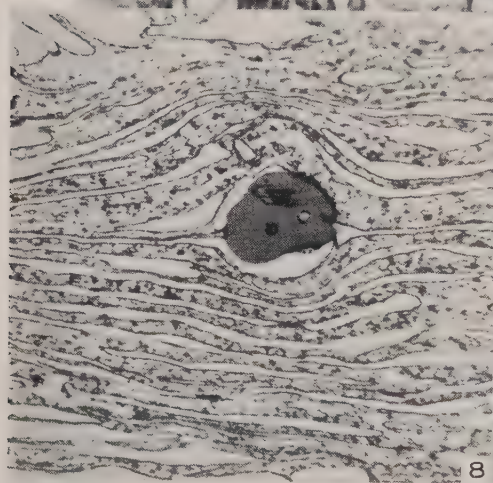




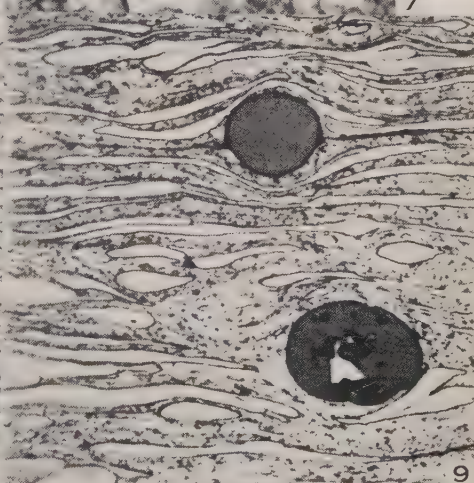




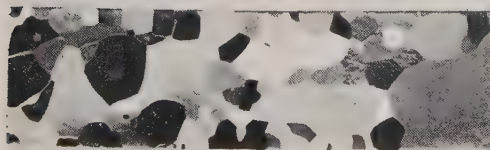
7



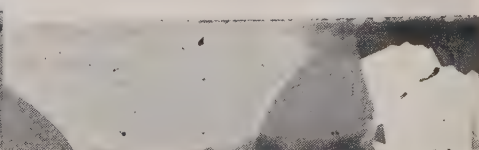
8



9

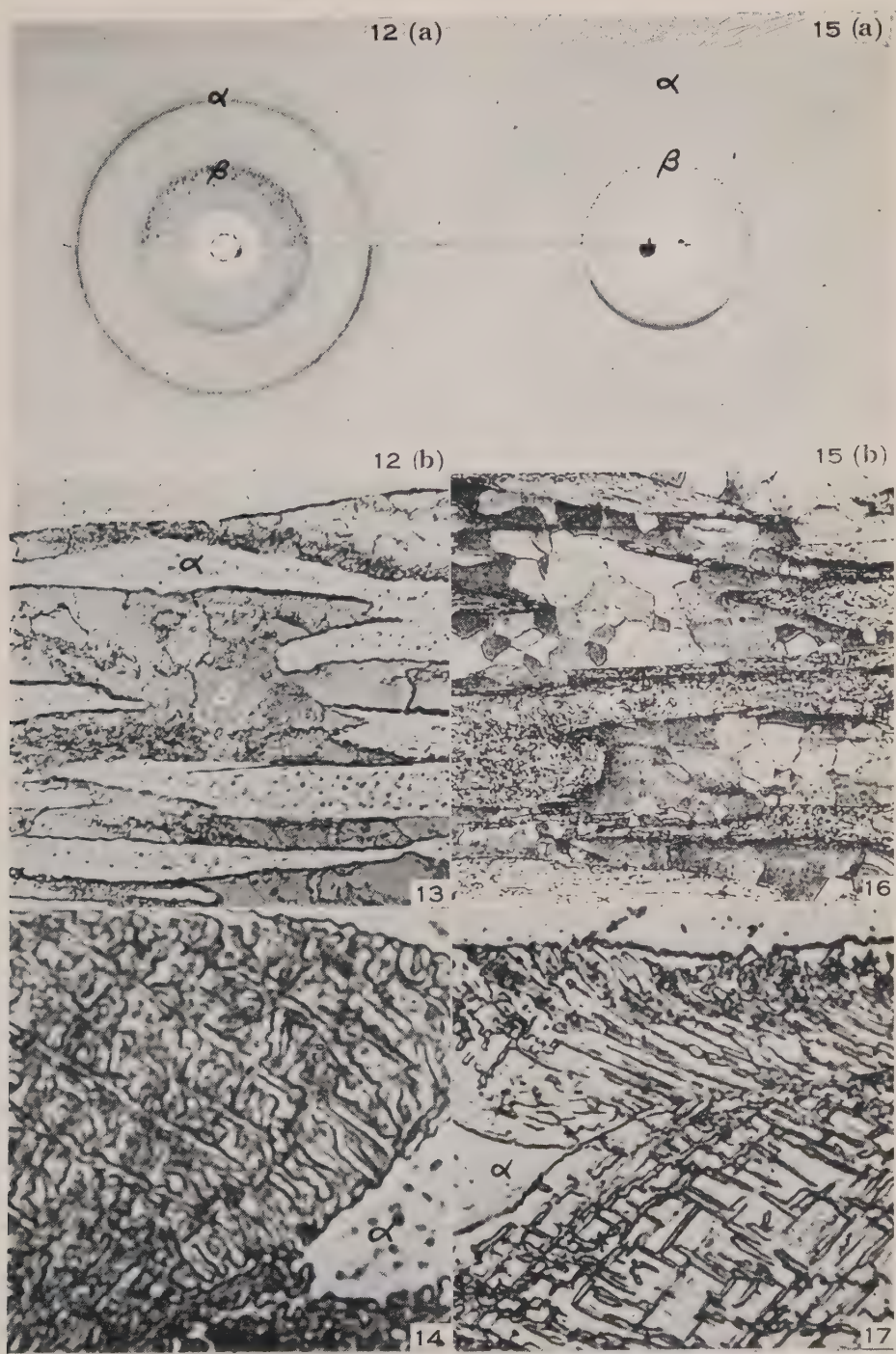


10



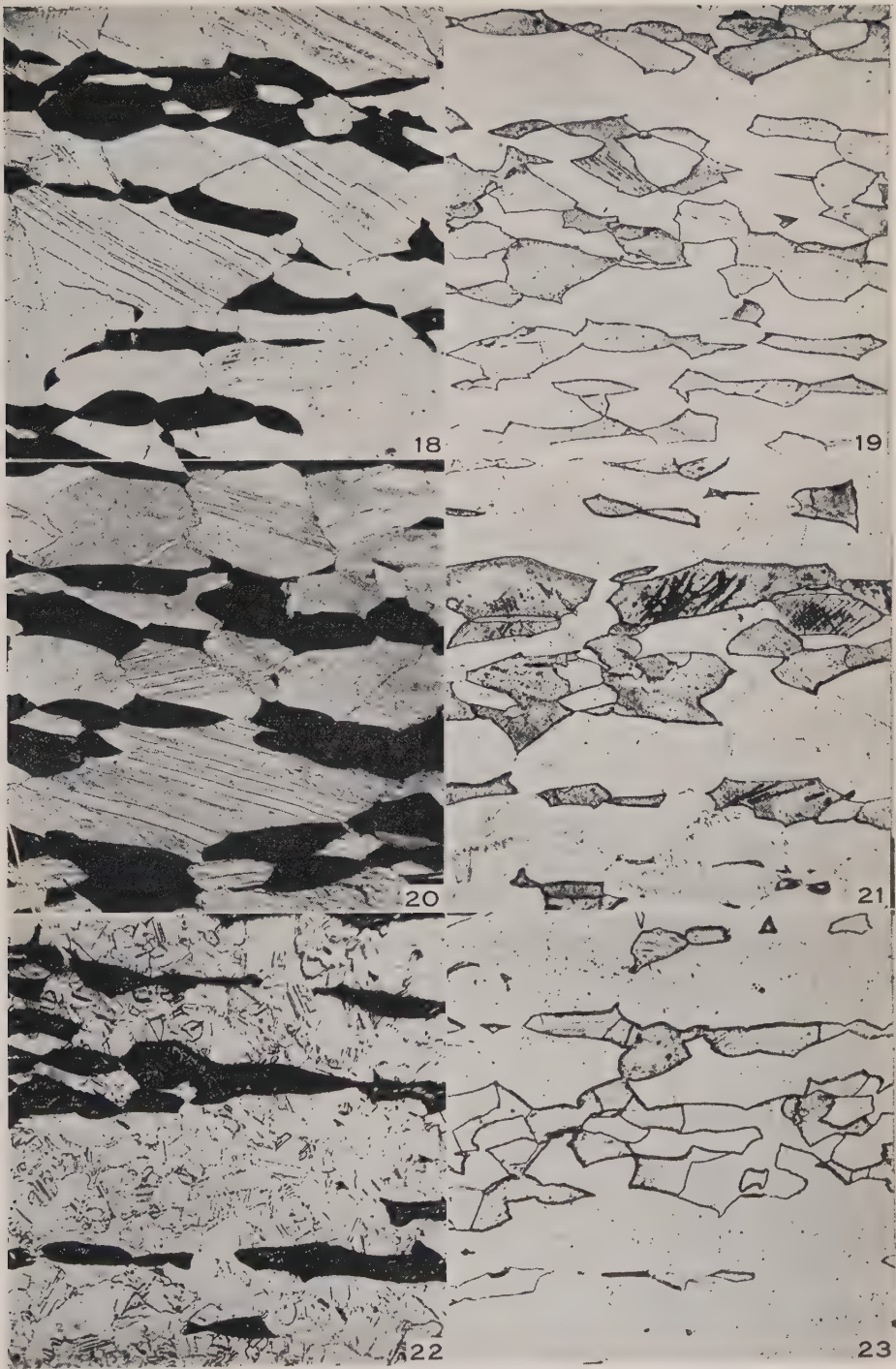
11





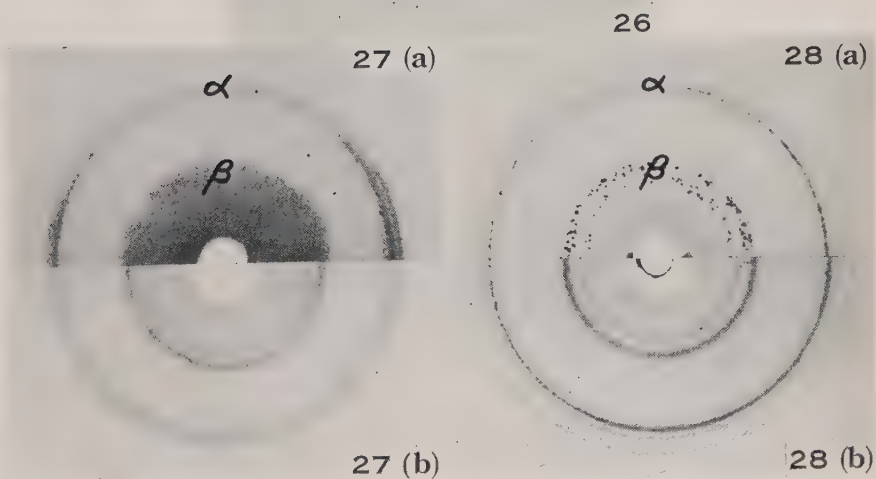
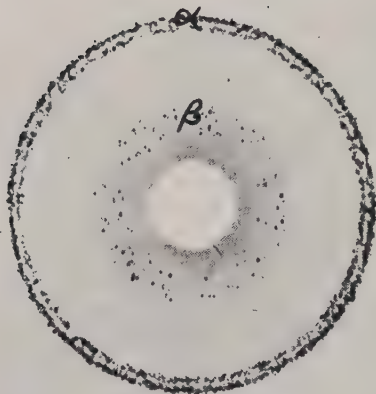
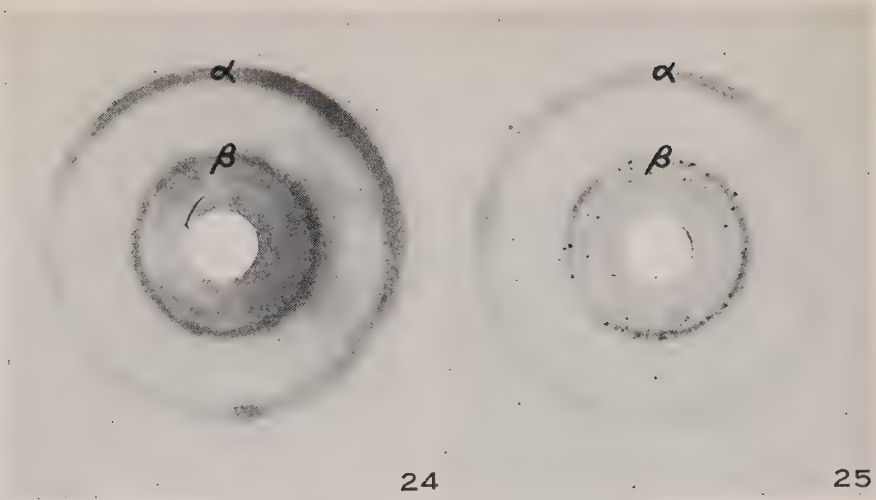






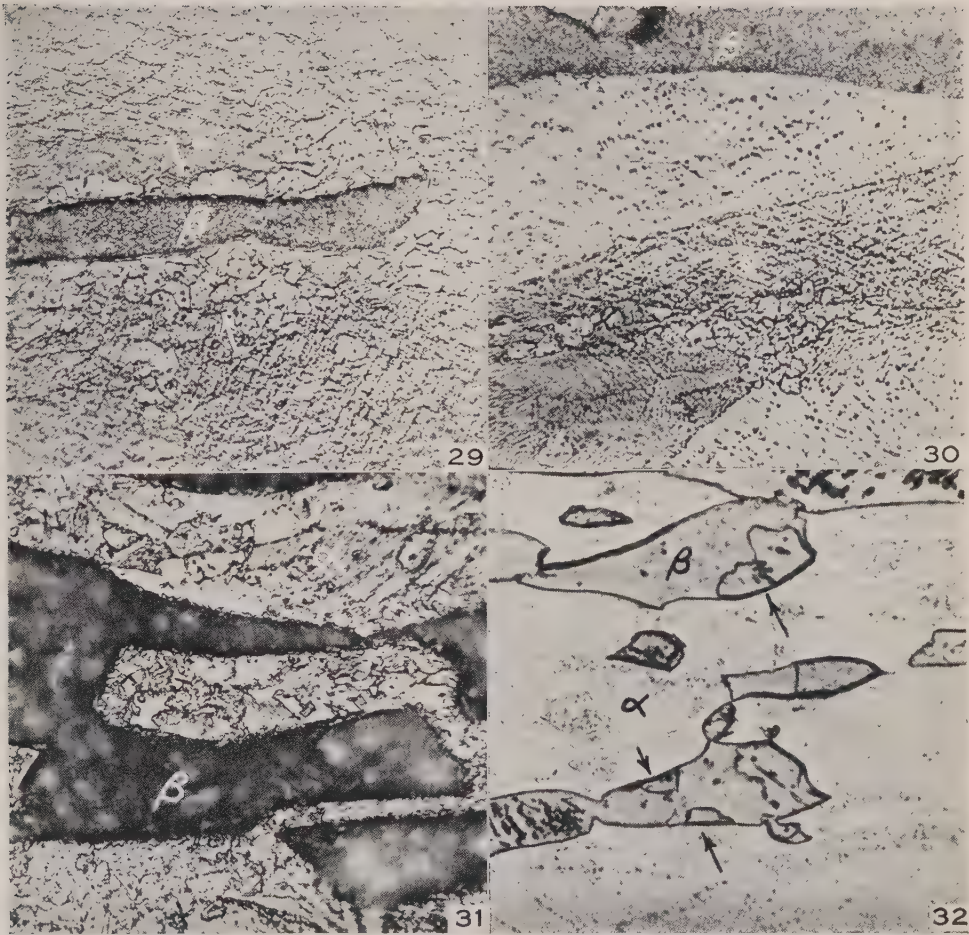
CLAREBROUGH.—DEFORMATION AND RECRYSTALLIZATION OF ALLOYS CONTAINING TWO PHASES











CLAREBROUGH.—DEFORMATION AND RECRYSTALLIZATION OF ALLOYS CONTAINING TWO PHASES



(iii) Comparison of temperatures at which recrystallization starts in  $\alpha$  and  $\beta$  phases

(1) Slowly cooled alloys

After low deformations :  $\alpha$  higher than  $\beta$ .

After high deformations :  $\alpha$  lower than  $\beta$ .

(2) Quenched alloys

All deformations :  $\alpha$  lower than  $\beta$ .

(i) (1) and (ii) (1) are explained by the influence of the state of order of the  $\alpha$  phase as discussed in Section III (b) (i). In the slowly cooled duplex alloys the  $\alpha$  phase is ordered although it contains slightly more than 25 atomic per cent. magnesium, whereas in the quenched alloys the  $\alpha$  phase is disordered. Thus similar behaviour to that shown by the single phase alloy  $\text{Ag}_3\text{Mg}$  is to be expected. The magnesium content of the  $\alpha$  phase in the quenched duplex alloy is higher than in the slowly cooled alloy, and this would tend to raise the recrystallization temperature of the  $\alpha$  in the quenched alloy. Since no such result was obtained after 47 per cent. reduction in area, the influence of composition must be small compared with the influence of the state of order.

(i) (2), (ii) (2), and (iii) (2) are explained by stress relief caused by precipitation. This is shown by the X-ray diagrams of the alloys annealed at 300 °C. after 80 per cent. reduction in area (Plate 5, Fig. 27). The Debye-Scherrer circle from the  $\beta$  phase of the quenched alloy is sharp, whereas that from the  $\beta$  of the slowly cooled alloy is diffuse. No recrystallization of the  $\beta$  has occurred in either case. Similar results are shown by the X-ray diagrams of the same alloys annealed at 375 °C. Here recrystallization of the  $\beta$  phase in the slowly cooled alloy is complete, but it has not started in the  $\beta$  of the quenched alloy where the Debye-Scherrer circle is sharp (Plate 5, Fig. 28). The type of precipitate occurring in these cases is identical with that shown before (Plate 3, Fig. 14).

The  $\beta$  phase in the quenched alloys contains less magnesium than in the slowly cooled alloys, so that a lower recrystallization temperature of the  $\beta$  in the quenched alloys would be expected. Since this is the reverse of the observed behaviour the influence of composition must, as for the  $\alpha$  phase, be small compared with the influence of precipitation.

The temperature at which recrystallization starts in the  $\beta$  phase of the quenched alloys is almost independent of deformation, since the effects of deformation and precipitation on recrystallization are opposite. The higher the deformation the greater is the rate of precipitation at any temperature, and with the greater amount of precipitate formed during the standard annealing time the stress relief before recrystallization is also greater.

The temperature at which recrystallization of the  $\beta$  is complete differs greatly in quenched and slowly cooled alloys because in quenched alloys recrystallization only goes to completion in 30 minutes when the amount of precipitate



formed is small, i.e. at high annealing temperatures. At these temperatures the effect of deformation becomes more marked, as shown by the decrease in the temperature for complete recrystallization of the  $\beta$  in quenched alloys with increasing deformation.

No explanation can be given for the fact (iii, 1) that in slowly cooled alloys increasing amounts of deformation lower the recrystallization temperature of the  $\alpha$  phase more than that of the  $\beta$ , so that a reversal of the order of recrystallization of the phases occurs. No such reversal occurs for quenched alloys because recrystallization of the  $\beta$  phase is affected by precipitation.

#### (d) *Mode of Recrystallization*

The mode of recrystallization of the  $\alpha$  phase varies with the phase proportions. When the volume proportion of  $\beta$  is small (approximately 10 per cent.), the  $\alpha$  phase nucleates preferentially at the  $\alpha$ - $\beta$  boundaries (Plate 6, Fig. 29). When the proportion of the  $\beta$  phase is increased to approximately 25 per cent., the  $\alpha$  phase tends to nucleate away from the  $\alpha$ - $\beta$  boundaries and closer to the centres of the  $\alpha$  areas (Plate 6, Fig. 30). When the proportion of the  $\beta$  phase is increased to approximately 50 per cent., nucleation in the  $\alpha$  occurs at random (Plate 6, Fig. 31). These results can be interpreted in terms of the deformation of the  $\alpha$  phase. When the proportion of the  $\beta$  phase is small, additional deformation of the  $\alpha$  occurs around the hard  $\beta$  particles as was illustrated for the copper-iron alloy in Plate 2, Figures 8 and 9. The presence of more  $\beta$  phase supports the  $\alpha$  crystals in the vicinity of the  $\alpha$ - $\beta$  boundaries, and the heaviest deformation occurs away from these boundaries. When the proportion of  $\beta$  is 50 per cent. or more, the  $\alpha$  crystals are surrounded by  $\beta$  crystals and heavy deformation cannot occur in the  $\alpha$  without corresponding heavy deformation of the  $\beta$ . Hence the deformation of the  $\alpha$  becomes more uniform, resulting in random nucleation.

For all phase proportions and for all deformations investigated, recrystallization of the  $\beta$  phase occurred by nucleation at the  $\alpha$ - $\beta$  boundary and growth of the new crystals inwards (Plate 6, Fig. 32). The same type of nucleation was also observed in the  $\beta$  crystals of duplex brass. This result indicates that the hard  $\beta$  crystals always deform to a greater extent at their boundaries with soft  $\alpha$  crystals and is thus in agreement with slip line observations (cf. Section III (a)). The mode of recrystallization of the  $\beta$  is not altered in the presence of precipitated  $\alpha$ . The  $\alpha$ - $\beta$  boundaries formed by precipitation do not act as centres for nucleation, as is common in age hardening alloys, indicating that here these regions are not centres of stress concentration.

### IV. DISCUSSION

#### (a) *Previous Experiments on Duplex Brass*

The results obtained with silver-magnesium alloys in the general case are similar to those obtained by Honeycombe and Boas for brass. The results for both alloy systems can be interpreted in the same way if all the contributing factors discussed in Section III (b) are taken into consideration.

The recrystallization temperature of the  $\beta$  phase in quenched silver-magnesium alloys is higher than in slowly cooled alloys, as stress relief is caused by precipitation. In quenched brass the influence of precipitation and of the order-disorder transformation in the  $\beta$  phase act together in the same direction and the observed increase in the recrystallization temperature of the  $\beta$  is therefore greater.

The results obtained by Honeycombe and Boas for a brass quenched from different temperatures in order to obtain different proportions of the phases show clearly the influence of precipitation. In their experiment, all the alloys were quenched from above the transformation temperature, and therefore the order-disorder transformation did not influence the results. They observed that the higher the quenching temperature, i.e. the greater the proportion of the  $\beta$  phase, the higher was the recrystallization temperature of this phase. The conclusion was drawn from this that the  $\beta$  phase was less deformed as its proportion increased. However, their results can now be explained more readily if stress relief by precipitation is taken into account. The degree of departure from equilibrium at the annealing temperatures increased as the quenching temperature was increased, and on annealing after deformation, precipitation of  $\alpha$  in  $\beta$  decreased the proportion of the  $\beta$  phase by as much as 50 per cent. The quantity of precipitate formed increased with the quenching temperature so that the stress relief increased correspondingly, resulting in an increase in the recrystallization temperature of the  $\beta$  phase.

They found also that after 86 per cent. reduction the  $\beta$  phase in an alloy stabilized at 300 °C. was fully recrystallized at 300 °C., whereas in an alloy quenched from 750 °C. the  $\beta$  phase did not start to recrystallize until 400 °C. and was not fully recrystallized until 500 °C. These results show the combined influence of precipitation and the order-disorder transformation. The behaviour is not due to the order-disorder transformation alone, as in an experiment designed to show this effect the  $\beta$  phase of the quenched alloy commenced to recrystallize at 300 °C. Precipitation in this case causes an increase in the recrystallization temperature of the  $\beta$  of 100 °C.

The present work has indicated that incorrect conclusions can be drawn regarding the deformation of the phases in a duplex alloy if the recrystallization temperatures of the phases are compared without due regard to the different heat treatments of the alloys before deformation. The order of recrystallization of the phases is a secondary result and it is their individual recrystallization temperatures and the change in these temperatures with deformation and heat treatment that is the primary consideration.

#### *(b) Influence of the Proportion of the Phases on their Relative Deformation*

The relation between the proportion of the phases and their relative deformation is indicated by the recrystallization results described in Section

III (b) (ii), where heat treatment and phase composition of the alloys are the same. The fundamental condition governing this relation is

$$p = p_{\alpha}v_{\alpha} + p_{\beta}v_{\beta}, \dots\dots\dots (1)$$

or

$$p = p_{\alpha} + v_{\beta}(p_{\beta} - p_{\alpha}), \dots\dots\dots (2)$$

where  $p$  = deformation per unit volume of the duplex alloy,

$p_{\alpha}$  = deformation per unit volume of the  $\alpha$  phase,

$p_{\beta}$  = deformation per unit volume of the  $\beta$  phase,

$v_{\alpha}$  = volume fraction of the  $\alpha$  phase,

$v_{\beta}$  = volume fraction of the  $\beta$  phase.

Thus if  $p_{\alpha}$  and  $p_{\beta}$  are constant over a range of values of  $v_{\beta}$  for the same  $p$ , then in this range

$$p_{\alpha} = p_{\beta} = p. \dots\dots\dots (3)$$

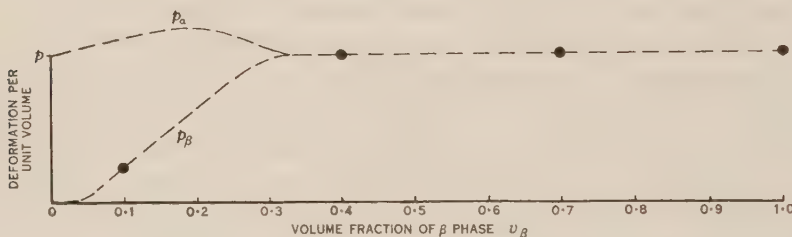


Fig. 4.—Influence of the volume fractions of the phases on their relative deformation.

The manner in which  $p_{\alpha}$  and  $p_{\beta}$  vary for the complete range of  $v_{\beta}$  is shown schematically in Figure 4. The recrystallization results have shown that  $p_{\alpha}$  and  $p_{\beta}$  are constant for alloys containing 40 per cent. and 70 per cent. of the  $\beta$  phase, so that in this range equation (3) must hold. Also, for

$$v_{\beta} = 1, \quad p_{\beta} = p,$$

and for

$$v_{\beta} = 0, \quad p_{\alpha} = p.$$

Both  $p_{\alpha}$  and  $p_{\beta}$  are plotted equal to  $p$  in the range of  $v_{\beta}$  from 0.7 to 1 as it seems unlikely that  $p_{\beta}$  will be less than  $p$  for  $v_{\beta}$  larger than 0.7. Owing to the brittleness of the silver-magnesium alloys containing more than 70 per cent. of the  $\beta$  phase, it was not possible to check this point experimentally.

The recrystallization results indicate that  $p_{\beta}$  is less in the alloy containing 10 per cent. of  $\beta$  than in that containing 40 per cent. of  $\beta$ , so that  $p_{\beta}$  is shown as lying below  $p$  for  $v_{\beta} = 0.1$ . For the purposes of the diagram it has been assumed that equation (3) ceases to apply when  $v_{\beta} = 0.3$ , because at this value the  $\alpha$  phase ceases to be a completely continuous matrix, and that  $p_{\beta}$  increases almost linearly with  $v_{\beta}$  in the range  $0 < v_{\beta} < 0.3$ . In this range  $p_{\beta} < p$  and hence  $p_{\alpha} > p$ . The maximum value of  $p_{\alpha}$  occurs when  $v_{\beta} = 0.2$ , and here  $p_{\alpha}$  is approximately 10 per cent. greater than  $p$ . This variation in  $p_{\alpha}$ , however, is not indicated clearly by the recrystallization results, as only alloys with  $v_{\beta}$  values of 0.1 and 0.4 were studied.



The deformation behaviour illustrated schematically by Figure 4 is a consequence of the phase distribution in the alloys. At low values of  $v_\beta$  the alloy consists of only a few isolated crystals of the hard  $\beta$  phase in a soft  $\alpha$  matrix. Thus the condition of the alloy is very similar to that of the copper-iron alloy shown in Plate 2, Figures 8 and 9. When  $v_\beta$  is very small most of the deformation occurs in the  $\alpha$  phase and for this reason the curve is shown approaching the  $v_\beta$  axis tangentially. There is, however, additional deformation in the  $\alpha$  phase around the hard  $\beta$  particles so that  $p_\alpha > p$ . When the volume proportion of the  $\beta$  phase exceeds approximately 30 per cent. the  $\alpha$  phase ceases to be a completely continuous matrix and the  $\alpha$  crystals start to become surrounded by hard  $\beta$  crystals. It is this condition of the alloy that prevents heavy deformations occurring in the  $\alpha$  phase without the  $\beta$  phase also deforming. This picture of the deformation process agrees with the mode of recrystallization described in Section III (d).

The fact that there are preferred zones of nucleation in the crystals indicates that the deformation is inhomogeneous owing to the interaction between crystals of different elastic limits at their boundaries. This effect can also be observed in single phase alloys, where the difference in elastic limits is caused by the crystal anisotropy. It has been found(7) that crystals with a small average extension ("hard" crystals) are more deformed near their boundaries with soft neighbours than at their centres, and that the soft crystals in turn are strengthened near their boundaries with hard crystals.

#### (c) *Precipitation and Stress Relief*

In silver-magnesium and copper-zinc alloys deformed after quenching the  $\beta$  phase is far removed from its equilibrium state, and on annealing, the  $\alpha$  phase precipitates rapidly in the form of flat sheets so that the strain energy associated with the precipitate is small(8).

For silver-magnesium alloys the specific volume of the  $\alpha$  phase is 18 per cent. less than that of the  $\beta$  phase, whereas for copper-zinc alloys this difference is only 2.5 per cent., suggesting that as the alloys behave similarly, the difference in specific volume does not play a very big part in the mechanism of stress relief. In both alloys precipitation of  $\alpha$  in  $\beta$  involves considerable diffusion which may remove part of the lattice distortion caused by deformation. The dislocation theory, however, suggests the most likely explanation of the mechanism of the stress relief. Stress relief involves a decrease in the density of dislocations in the deformed  $\beta$  and this decrease could be brought about by diffusion of the dislocations to the zones of misfit at the new  $\alpha$ - $\beta$  boundaries created by precipitation. Since the precipitate is very fine, the distances over which diffusion would have to occur are small and in the cases studied are of the order of  $2.5 \times 10^{-5}$  cm. A similar type of explanation has been advanced by Cahn(9) for stress relief by polygonization.

#### (d) *Order-Disorder and Recovery*

The alloys used in this work provide examples of superlattices of different stabilities. The  $\beta$  phase in the silver-magnesium alloys is ordered at all temperatures up to the melting point and order can only be destroyed by cold work.



The order-disorder transformation in the  $\alpha$  phase, however, can be suppressed by quenching. In  $\beta$  brass it is not possible to suppress completely the order-disorder transformation by quenching, but it is likely that the degree of order is less in the quenched than in the slowly cooled alloy(5). In all these alloys, regardless of whether disorder is produced by cold work or by quenching, reordering of the deformed alloys occurs before recrystallization and may cause recovery.

The recrystallization behaviour of  $\text{Ag}_3\text{Mg}$  and  $\beta$  brass described in Section III (b) (ii) and the small decrease in recrystallization temperature of the  $\beta$  phase in duplex silver-magnesium alloys with increasing deformation (Section III (c)) may be explained if it is accepted that reordering can reduce lattice strains and thus act as a recovery process. Further work on the influence of reordering on recovery and on recrystallization is planned.

## V. ACKNOWLEDGMENTS

The author wishes to thank Dr. W. Boas for suggesting the problem and for much helpful advice and discussion during the course of the work. Thanks are also due to Professor E. J. Hartung for laboratory facilities in the Chemistry School, University of Melbourne, to the staff of the Defence Research Laboratories for the analyses, and to Mr. G. R. Perger and Mr. J. J. Batten, Division of Tribophysics, for assistance with the experimental work.

## VI. REFERENCES

- (1) HONEYCOMBE, R. W. K., and BOAS, W.—*Aust. J. Sci. Res. A* **1**: 70 (1948).
- (2) ANDREWS, K. W., and HUME-ROTHERY, W.—*J. Inst. Met.* **69**: 485 (1943).
- (3) LETNER, H. R., and SIDHU, S. S.—*J. Appl. Phys.* **18**: 833 (1947).
- (4) UNCKEL, H.—*J. Inst. Met.* **61**: 171 (1937).
- (5) SMITH, C. S.—*Trans. Amer. Inst. Min. (Metall.) Engrs.* **152**: 144 (1943).
- (6) CLAREBROUGH, L. M.—*Nature* **165**: 39 (1950).
- (7) BOAS, W., and HARGREAVES, M. E.—*Proc. Roy. Soc. A* **193**: 89 (1948).
- (8) NABARRO, F. R. N.—*Proc. Roy. Soc. A* **175**: 519 (1940).
- (9) CAHN, R. W.—*Phys. Soc. Rep. Conf. on Strength of Solids*, p. 136 (1948).

## EXPLANATION OF PLATES 1-6

### PLATE 1

- Fig. 1.—Polished and etched surface of duplex silver-magnesium alloy after 10 per cent. compression. The  $\alpha$  phase shows heavy slip but there is no slip in the  $\beta$  phase.  $\times 460$ .
- Fig. 2.—Same specimen as shown in Figure 1. Slip lines from the  $\alpha$  crystals continue for a short distance into the  $\beta$  crystals without change in direction.  $\times 460$ .
- Fig. 3.—Duplex silver-magnesium alloy given 60 per cent. reduction by rolling, polished and etched, and then given 10 per cent. compression. Slip lines are present in both phases but the  $\alpha$  phase appears more heavily deformed than the  $\beta$ .  $\times 460$ .
- Fig. 4.—Duplex silver-magnesium alloy given 70 per cent. reduction by rolling, polished and etched, and then given 10 per cent. compression. Heavy slip bands are present in the  $\beta$  phase.  $\times 460$ .
- Fig. 5.—Polished and etched surface of duplex silver-magnesium alloy containing 25 volume per cent. of the  $\beta$  phase after 5 per cent. compression. The  $\alpha$  shows marked slip but the  $\beta$  none.  $\times 400$ .

Fig. 6.—Polished and etched surface of duplex silver-magnesium alloy containing 70 volume per cent. of the  $\beta$  phase after 5 per cent. compression. The  $\alpha$  and  $\beta$  phases show marked slip.  $\times 400$ .

## PLATE 2

Fig. 7.—Structure of copper-iron alloy as cast. The inclusion in the centre of the field is a particle of iron oxide, the white areas are dendrites of iron, and the matrix is copper containing a precipitate of iron.  $\times 500$ .

Fig. 8.—Copper-iron alloy after 90 per cent. reduction by rolling. The flow in the matrix is disturbed around the particle of iron oxide.  $\times 500$ .

Fig. 9.—Another field of the same specimen as in Figure 8.  $\times 500$ .

Fig. 10.— $\beta$  brass slowly cooled from 600 °C. to room temperature, deformed to 20 per cent. reduction by rolling, and annealed for 30 minutes at 450 °C. The specimen has almost fully recrystallized.  $\times 6.5$ .

Fig. 11.— $\beta$  brass quenched from 600 °C., deformed to 20 per cent. reduction by rolling, and annealed for 30 minutes at 450 °C. Recrystallization has just started.  $\times 6.5$ .

## PLATE 3

Fig. 12.—Back reflection X-ray diagrams of duplex silver-magnesium alloys annealed for 30 minutes at 350 °C. after 80 per cent. reduction in area.

(a) Slowly cooled alloy.  $\beta$  almost fully recrystallized,  $\alpha$  fully.

(b) Quenched alloy. No recrystallization of  $\beta$ .  $\alpha$  fully.

The Debye-Scherrer circles of the two phases are marked  $\alpha$  and  $\beta$ .

Fig. 13.— $\beta$  phase in slowly cooled silver-magnesium alloy almost fully recrystallized.  $\times 1060$ .

Fig. 14.—Widmanstätten structure formed by precipitation of  $\alpha$  in the  $\beta$  phase of the quenched silver-magnesium alloy. No recrystallization of  $\beta$ .  $\times 3250$ .

Fig. 15.—Back reflection X-ray diagrams of duplex copper-zinc alloys annealed for 30 minutes at 300 °C. after 80 per cent. reduction in area.

(a) Slowly cooled alloy.  $\beta$  fully recrystallized.

(b) Quenched alloy. No recrystallization of  $\beta$ .

Fig. 16.— $\beta$  phase in slowly cooled copper-zinc alloy fully recrystallized.  $\times 490$ .

Fig. 17.—Widmanstätten structure formed by precipitation of  $\alpha$  in the  $\beta$  phase of the quenched copper-zinc alloy. No recrystallization of  $\beta$ .  $\times 2030$ .

## PLATE 4

Photomicrographs ( $\times 245$ ) of duplex silver-magnesium alloys slowly cooled from 600 °C., deformed to 34 per cent. reduction in area, and annealed for 30 minutes at different temperatures.

Fig. 18.—Annealed at 325 °C. No recrystallization of  $\alpha$ .

Fig. 19.—Same specimen as in Figure 18. No recrystallization of  $\beta$ .

Fig. 20.—Annealed at 350 °C. No recrystallization of  $\alpha$ .

Fig. 21.—Same specimen as in Figure 20.  $\beta$  started to recrystallize.

Fig. 22.—Annealed at 400 °C.  $\alpha$  fully recrystallized.

Fig. 23.—Same specimen as in Figure 22.  $\beta$  fully recrystallized.

## PLATE 5

Back reflection X-ray diagrams of silver-magnesium alloys.

Fig. 24.—Slowly cooled alloy annealed at 325 °C. after 34 per cent. reduction. No recrystallization of  $\beta$  or  $\alpha$ .

Fig. 25.—Slowly cooled alloy annealed at 350 °C. after 34 per cent. reduction.  $\beta$  started to recrystallize,  $\alpha$  none.

Fig. 26.—Slowly cooled alloy annealed at 400 °C. after 34 per cent. reduction.  $\beta$  and  $\alpha$  fully recrystallized.

Fig. 27.—Slowly cooled and quenched alloys annealed at 300 °C. after 80 per cent. reduction in area.

(a) Slowly cooled alloy. No recrystallization of  $\beta$ . Debye-Scherrer circle for  $\beta$  is diffuse.

(b) Quenched alloy. No recrystallization of  $\beta$ . Debye-Scherrer circle for  $\beta$  is sharp.

Fig. 28.—Slowly cooled and quenched alloys annealed at 375 °C. after 80 per cent. reduction in area.

(a) Slowly cooled alloy.  $\beta$  fully recrystallized.

(b) Quenched alloy. No recrystallization of  $\beta$ .

## PLATE 6

The mode of recrystallization of the phases in silver-magnesium alloys. Arrows indicate new grains.

Fig. 29.—Recrystallization of the  $\alpha$  phase starts at the  $\alpha$ - $\beta$  boundary when the volume fraction of the  $\beta$  is 0.1.  $\times 1080$ .

Fig. 30.—Recrystallization of the  $\alpha$  phase starts away from the  $\alpha$ - $\beta$  boundaries when the volume fraction of the  $\beta$  is 0.25.  $\times 1080$ .

Fig. 31.—Recrystallization of the  $\alpha$  phase starts at random when the volume fraction of the  $\beta$  is 0.5.  $\times 1080$ .

Fig. 32.—Recrystallization of the  $\beta$  phase always starts at the  $\alpha$ - $\beta$  boundary.  $\times 500$ .

# SELF-DIFFUSION IN TIN CRYSTALS

By P. J. FENSHAM\*

[Manuscript received October 25, 1949]

## Summary

The rate of self-diffusion in single crystals of white tin of various orientations has been measured using the radio-active isotope  $\text{Sn}^{113}$ . The ratio of the diffusion coefficients parallel and perpendicular to the tetragonal (*c*) axis is approximately 2 at 180 °C. and approximately 3 at 223 °C. The diffusion coefficients in both directions at various temperatures are given by the Arrhenius expressions  $D_c = 1 \cdot 2 \times 10^{-5} e^{-10,500/RT}$  cm.<sup>2</sup>sec.<sup>-1</sup> and  $D_a = 3 \cdot 7 \times 10^{-8} e^{-5,900/RT}$  cm.<sup>2</sup>sec.<sup>-1</sup>. The anisotropy is discussed in terms of the vacancy mechanism of diffusion.

## I. INTRODUCTION

Self-diffusion in single crystals is the simplest case of migration of atoms in solids and is important both from the theoretical aspect of the diffusion mechanism and with regard to reactions in solids. The measurement of the rate of self-diffusion is possible only by the use of radio-active isotopes as tracer atoms, and whereas initially such measurements were restricted to elements with naturally occurring isotopes (lead and bismuth), the technique can now be applied to most elements. Up to date, measurements have been made on polycrystalline Pb(1, 2), Bi(3), Cu(4, 5), Au(6), Ag(7), Zn(8, 9), and with Pb, Bi, Cu, and Zn also in single crystals.

Single crystals were selected for our experiments as the diffusion rate along grain boundaries is usually different from that within the crystal.† Metallic tin was chosen because, owing to the tetragonal structure, the rate of diffusion was expected to be anisotropic, and it was hoped that its relation with other anisotropic properties would become apparent.

The first measurements of the rates of self-diffusion in different directions were made by Seith(3) on bismuth crystals. Seith used two sets of crystals, measuring in a direction either parallel or perpendicular to the principal axis. His results showed that, near the melting point, the rate perpendicular to this axis was  $10^6$  times that parallel to the axis.

The rate of self-diffusion in Zn crystals has been investigated by Miller and Banks(8, 9). They measured the rate parallel to the hexagonal axis directly, but calculated that perpendicular to this axis from measurements in a direction inclined at an angle  $\theta$  to the axis, using the equation

$$\ln D_\theta = \cos^2 \theta \ln D_c + \sin^2 \theta \ln D_a, \dots\dots\dots (1)$$

\* Chemistry Department, University of Melbourne, and Division of Tribophysics, C.S.I.R.O.

† This effect is discussed in a subsequent paper.



where  $D_\theta$  is the diffusion coefficient in the measured direction and  $D_c$  and  $D_a$  are the values along the respective axes. It is usual to express the effect of temperature on the diffusion coefficient by the Arrhenius equation

$$D = Ae^{-Q/RT}, \dots\dots\dots (2)$$

where  $A$  is a temperature independent factor and  $Q$  the apparent activation energy for diffusion, the significance of which will be discussed later in this paper. Since such an equation holds for both  $D_c$  and  $D_a$  there are four constants which have to be determined for zinc. Values of these constants have been obtained by Miller and Banks from a number of measurements of  $D_\theta$  made at various temperatures and at various orientations.

There are two objections against their procedure. Firstly, the equation (1) is not correct.\* As shown in Appendix I, the equation giving the diffusion coefficient in a direction  $\theta$  is, for hexagonal and tetragonal crystals,

$$D_\theta = D_c \cos^2 \theta + D_a \sin^2 \theta, \dots\dots\dots (3)$$

or

$$D_\theta = (D_a - D_c) \sin^2 \theta + D_c, \dots\dots\dots (3a)$$

or

$$D_\theta = A_c e^{-Q_c/RT} \cos^2 \theta + A_a e^{-Q_a/RT} \sin^2 \theta \dots\dots\dots (3b)$$

Secondly if, as in Miller and Banks's experiments, temperature and orientation are altered simultaneously, the four constants have to be found by solving (3b) from the measured values of  $D_\theta$  using the method of least squares. For this purpose, an arbitrary set of four  $D_\theta$  values is used to calculate preliminary estimates of the  $A$  and  $Q$ . These are then corrected to give a best fit to all the observations. Equation (3b) is a non-linear equation, and on applying this method to the available results the final values obtained are found to depend markedly on the initial selection of the four  $D_\theta$  values. The variations are sufficiently large to preclude this method of solution. A change in the design of the experiment is therefore essential.

The procedure adopted in our investigation(10) was to change only one of the variables, temperature or orientation, in each set of experiments. In crystals of various orientations, the diffusion rates were measured at a constant temperature and thus  $D_c$  and  $D_a$  were calculated at this temperature. By carrying out similar sets of experiments at other temperatures, the values of  $A_c$ ,  $A_a$ ,  $Q_c$ , and  $Q_a$  were then obtained.

## II. EXPERIMENTAL

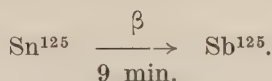
Briefly the experimental procedure was as follows. A thin layer containing the radio-active isotope was plated on a flat face of a single crystal. After diffusion was allowed to occur the concentration at various depths was found by cutting thin layers parallel to the diffusion interface and measuring their activity.

\* Recalculation of their results using equation (3) gave the following values :

$$\begin{array}{ll} Q_c = 1.93 \times 10^4 \text{ cal./g. atom} & A_c = 0.4 \times 10^3 \text{ cm.}^2 \text{ day}^{-1} \\ Q_a = 3.2 \times 10^4 \text{ cal./g. atom} & A_a = 13.8 \times 10^6 \text{ cm.}^2 \text{ day}^{-1}. \end{array}$$

*(a) Separation of Radio-Active Materials*

Radio-active tin was obtained from the United States Atomic Energy Commission. The material supplied was a mixture of radio-active  $\text{Sn}^{113}$  and  $\text{Sb}^{125}$ , together with a large quantity of inactive tin. The mixed isotopes, as chlorides, were in solution in 3N HCl. They had been prepared by irradiating an inactive tin target with neutrons in the uranium chain reacting pile. The  $\text{Sb}^{125}$  was formed as a decay product of a short-lived tin isotope,  $\text{Sn}^{125}$ , which had also been produced in the pile,



The properties of the isotopes are set out in Table 1.

TABLE 1  
PROPERTIES OF ISOTOPES AS RECEIVED

Isotope	Half-life	Radiation	Specific Activity
$\text{Sn}^{113}$	100 days	$K, \gamma, e^-$	0.0138 mc./ml.
$\text{Sb}^{125}$	2.7 years	$\beta^-, \gamma$	0.0060 mc./ml.

Since the presence of even minute quantities of  $\text{Sb}^{125}$  would give errors in the counts of tin activity in any sample, it was essential to remove this isotope completely from the mixture.

The separation of the antimony from the tin was carried out by successive precipitations from solution by adding metallic tin. As only negligible gravimetric amounts of  $\text{Sb}^{125}$  were present, inactive antimony was added to act as carrier in the precipitation of the active isotope. For the first precipitation, 0.5 g. per mc. of inactive antimony chloride was added. Pure tin, just in excess of that necessary to precipitate the antimony completely, was placed in the 3N HCl solution. The black precipitate of metallic antimony was then removed from the solution and the precipitation repeated after adding a second smaller quantity of carrier antimony. The activity of the antimony precipitate was measured after each precipitation by means of a Geiger counter. The precipitations were continued until the activity of the precipitate had dropped to such a value that the amount of radio-active  $\text{Sb}^{125}$  remaining in the solution made only a negligible contribution to the total activity.

The hydrochloric acid solution of the tin isotopes was then used in the plating bath for the tin crystals.

*(b) Preparation of Tin Crystals*

The tin used for the preparation of the crystals was examined spectrographically and found to have a purity of at least 99.998 per cent. The maximum impurity was 0.0005 per cent. Cu.

Single crystals were grown by a method similar to that described by Hasler(11), the moulds being of rectangular cross-section with dimensions 9 by 1.3 by 0.9 cm. The specimens for the diffusion experiments were obtained from the large crystals by cutting off 1.5 cm. lengths with a fine jeweller's saw. It was important to remove completely the deformation produced during the cutting to ensure that no recrystallization would occur during the experiment. This was done by mechanical polishing followed by electropolishing to below the deformed layer. The solution(12) used for electropolishing is given in Table 5, Appendix II.

After the crystals had been annealed at 150 °C. for several hours, their crystallographic orientation was determined by the method of etch pits using a two-circle goniometer(13). The etching conditions are given in Table 6, Appendix II.

The diffusion face of the specimen was then prepared as flat as possible by carefully polishing it on emery paper followed by a light electropolish. This was repeated several times until a flat undeformed surface was obtained, except for the slight rounding over of the edges, which was unavoidable in the polishing method. For the diffusion experiment a thin layer of the active-inactive tin mixture was plated on to the prepared diffusion interface by the method given in Table 7, Appendix II.

In this way an adherent thin layer (0.02–0.03 mm.) of tin was obtained which had an activity of about 16,000 counts per minute.

### (c) Diffusion-Penetration Measurements

The specimens were sealed off in a Pyrex tube at  $10^{-3}$  mm. Hg pressure and placed in an oil-bath at the desired temperature of the diffusion for periods of 1–2 days. The temperature was controlled to  $\pm 0.3$  °C.

After the diffusion, the sides of the crystal parallel to the diffusion direction were removed to a depth of 0.5 mm. to eliminate the serious effect of surface diffusion and the edge effects introduced by the polishing. The specimen was then sectioned parallel to the diffusion interface by means of a microtome. Eight to ten sections ( $1-2 \times 10^{-2}$  mm. thick) were taken and their exact thickness obtained from their weight and cross-sectional area. Finally, the remaining crystal was etched and examined to ensure that no recrystallization had occurred during the heating. Only one specimen had recrystallized and this was not used in subsequent calculations.

The diffusion which had taken place was followed by measuring the concentration of radio-active tin in each section.

$\text{Sn}^{113}$  decays(14) by  $K$  capture to active  $\text{In}^{113*}$  which in turn gives inactive  $\text{In}^{113}$  according to



The amount of indium ( $< 10^{-6}$  per cent.) present in the plated layer is far below the sensitivity of the spectrographic analysis and negligible relative to the impurities present in the tin.



The main radiations emitted during the decay are  $\gamma$ -rays of energies 85 and 390 keV.

The radio-activity of a sample was measured by means of a General Electric GM4 metal counter tube, which was connected to a scale of 16 counting unit and an electromechanical recorder. The background of the counter was about 20 counts per min.

In order to eliminate decay effects and any counting fluctuations that might occur during a set of readings, all counts were referred to a standard specimen, which was a mixture of the active and inactive tin isotopes (Sb free). In measuring the activity of the diffusion sections it was important that the sections should be placed in the same geometrical position relative to the counter. To ensure this they were dissolved in dilute nitric acid and the tin precipitated as gelatinous tin hydroxide on small flat aluminium dishes (2.5 cm. diameter). This procedure also eliminated the need to correct for self-absorption of the radiation, since, for the energies involved, the absorption was negligible in the thin layer of tin hydroxide.

The concentration of each section was expressed in the arbitrary units of counts per min. per mg.

#### (d) Accuracy of Measurements

The sections were weighed to  $\pm 0.1$  mg. and this gave an accuracy of 1 per cent. in weight which, with 0.5 per cent. in the area, gave a probably accuracy of 1.5 per cent. in the thickness of the sections.

The probable error in the activity of a section is given by the square root of the number of counts. The activities varied from 1600 to 48 counts per min. In general, sufficient counts were taken from each section to give an accuracy of 1 per cent. For a few sections of low activity the accuracy was only 2 per cent. and these points were weighted accordingly in the calculations. The very small counts of still deeper sections were not used since they were of the same order as the background and the counting times required to obtain accurate results became prohibitive. The error in the time is negligible since this was accurate to at least  $\pm 1$  min. which represents an error of less than 0.1 per cent. The errors in the other measured quantities were  $\pm 1^\circ$  in the orientation and  $\pm 0.3^\circ\text{C}$ . in the temperature of diffusion.

### III. EVALUATION OF RESULTS

In our experiment the diffusion can be considered as linear. The boundary conditions for which the diffusion equation (see Appendix I, equation (13)) has to be solved are  $t=0$ ;  $c=c_0$  (const.) for  $0 \leq x \leq h$ ;  $c=0$  for  $h < x \leq l$ ;  $h/l < 1$ , and the solution (15) is

$$c = \frac{c_0}{\sqrt{\pi D_0 t}} e^{-x^2/4D_0 t},$$

where  $c$  is the concentration,

$t$  is the time,

$x$  is the diffusion coordinate,

$h$  is the thickness of the plated layer, and

$l$  is the specimen length.



In order to obtain a value of  $D_0$  for a particular crystal,  $\ln c$  is plotted against  $\bar{x}^2$  where  $\bar{x}$  is the mean distance of a section from  $x=0$ . Some typical graphs of this type are shown in Figure 1.

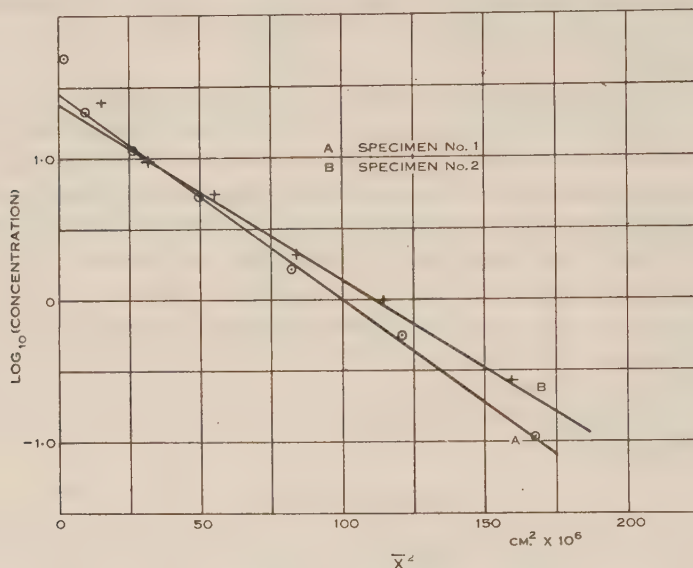


Fig. 1.—Diffusion penetration curves for specimens 1 and 2 (see Table 2).

From the slope ( $-1/4D_0t$ ) of the straight line, the value of  $D_0$  can be obtained. In all cases the first point was found to be high and was neglected. The justification for this procedure will be discussed later. The diffusion distances are

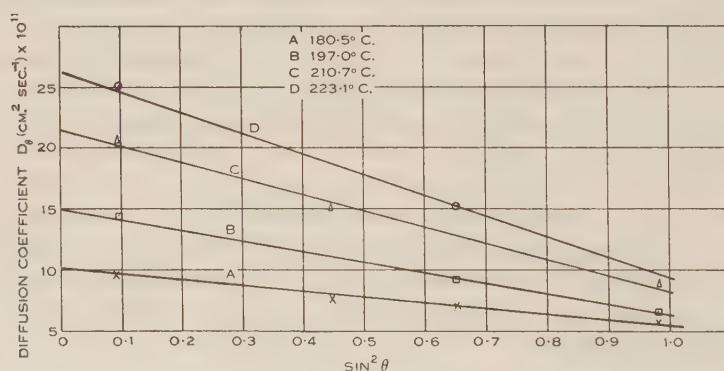


Fig. 2.—Diffusion coefficient as a function of the orientation at various temperatures.

measured at room temperature, while the actual diffusion occurs at elevated temperatures, and it is therefore necessary to correct the measured distances for thermal expansion. The coefficients of expansion given by Bridgman(16) were used to make this correction in  $D_0$ .

The results obtained for crystals of different orientations and at various temperatures of diffusion are summarized in Table 2.

TABLE 2  
DIFFUSION COEFFICIENTS

Specimen No.	Orientation ( $\theta^\circ$ )	Temperature ( $t^\circ\text{C.}$ )	$1/T^\circ\text{K.}$	Diffusion Coefficient $D_\theta$ ( $\text{cm.}^2\text{sec.}^{-1}$ )
9	18	180.5	$2.211 \times 10^{-3}$	$9.76 \times 10^{-11}$
23	42	180.5	$2.211 \times 10^{-3}$	$7.63 \times 10^{-11}$
16	54	180.5	$2.211 \times 10^{-3}$	$7.11 \times 10^{-11}$
1	82	180.5	$2.211 \times 10^{-3}$	$5.67 \times 10^{-11}$
11	18	197.0	$2.126 \times 10^{-3}$	$14.3 \times 10^{-11}$
17	54	197.0	$2.126 \times 10^{-3}$	$9.13 \times 10^{-11}$
3	82	197.0	$2.126 \times 10^{-3}$	$6.69 \times 10^{-11}$
12	18	210.7	$2.076 \times 10^{-3}$	$20.6 \times 10^{-11}$
21	42	210.7	$2.076 \times 10^{-3}$	$15.2 \times 10^{-11}$
2	82	210.7	$2.076 \times 10^{-3}$	$8.80 \times 10^{-11}$
13	18	223.1	$2.015 \times 10^{-3}$	$25.1 \times 10^{-11}$
15	54	223.1	$2.015 \times 10^{-3}$	$15.1 \times 10^{-11}$
18	54	223.1	$2.015 \times 10^{-3}$	$14.7 \times 10^{-11}$
25	75	223.1	$2.015 \times 10^{-3}$	$10.9 \times 10^{-11}$

According to equation (3a) each set of  $D_\theta$  values at constant temperature plotted against  $\sin^2 \theta$  should lie on a straight line. Figure 2 shows that this is so.

From these results the values of  $D_c$  and  $D_a$  at each temperature were obtained (see Table 3).

TABLE 3  
PRINCIPAL DIFFUSION COEFFICIENTS

Temperature ( $t^\circ\text{C.}$ )	Diffusion Coefficient ( $\text{cm.}^2 \text{sec.}^{-1}$ )	
	$D_c$	$D_a$
180.5	$9.89 \times 10^{-11}$	$5.47 \times 10^{-11}$
197.0	$15.05 \times 10^{-11}$	$6.67 \times 10^{-11}$
210.7	$21.44 \times 10^{-11}$	$8.10 \times 10^{-11}$
223.1	$26.84 \times 10^{-11}$	$9.64 \times 10^{-11}$

If these principal diffusion coefficients are assumed to have temperature dependence of the Arrhenius type, viz.

$$\left. \begin{aligned} D_c &= A_c e^{-Q_c/RT}, \\ D_a &= A_a e^{-Q_a/RT}, \end{aligned} \right\} \dots\dots\dots (4)$$

it is possible to calculate the values of the four independent constants by determining the slope and axial intercept of the linear plot of  $\ln D$  v.  $1/T^\circ$  (see Fig. 3). The results together with the probable errors are shown in Table 4.

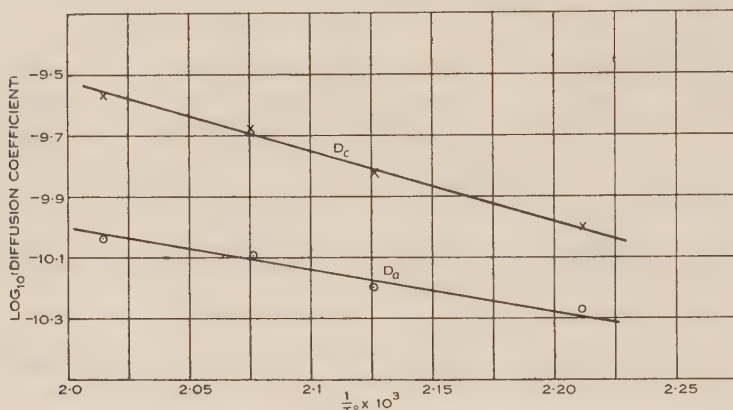


Fig. 3.—Logarithms of  $D_c$  and  $D_a$  as a function of temperature.

TABLE 4  
ARRHENIUS CONSTANTS FOR TIN

Constant			Parallel to	
			a-Axis	c-Axis
$A$ (cm. <sup>2</sup> sec. <sup>-1</sup> )	..	..	$3.7 \pm 1.5 \times 10^{-8}$	$1.2 \pm 0.5 \times 10^{-5}$
$Q$ (cal./g. atom)	..	..	$5.9 \pm 0.4 \times 10^3$	$10.5 \pm 0.5 \times 10^3$

The method of least squares was used for each step of the calculations(17). The error in the values of  $D_0$  was found to be approximately 4 per cent.

It can be seen that both  $A$  and  $Q$  depend on the crystallographic direction and that the high value of  $Q$  is associated with the high value of  $A$ . A similar result was also found with zinc(9) and bismuth(3). It is interesting to note that functional relationships between the frequency factor and the activation energy have been observed with other rate processes, e.g. fluidity(18) and chemical reaction rates(19).

## IV. DISCUSSION

(a) *Anisotropy*

The diffusion coefficient,  $D$ , is defined as the amount of substance diffusing in unit time across unit area through a unit concentration gradient. Hence the magnitude of the diffusion coefficient gives a direct measure of the diffusion rate.

It is evident from the results in Table 3 that the rate of self-diffusion in tin is anisotropic. The ratio of the rates,  $D_c/D_a$ , is approximately 3 at 220 °C. and 2 at 180 °C. Extrapolation to lower temperatures indicates that the rates would be equal at about 125 °C. and below this temperature  $D_a$  would be greater than  $D_c$ . However, whether such an extrapolation is justifiable is not certain, since no experiments were carried out in this region. It has been noted previously that if the equations (4) are valid both the activation energy,  $Q$ , and the factor,  $A$ , depend considerably on the crystallographic direction.

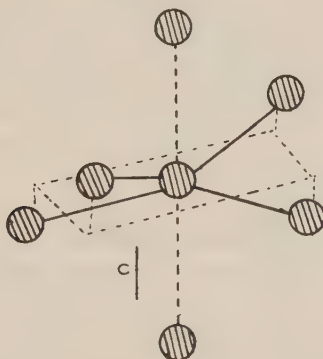


Fig. 4.—The neighbours of an atom in the structure of white tin.

An attempt has been made by Huntington(20) to account for the difference in the  $A$  values for zinc in terms of the geometry of the lattice. Such a qualitative consideration may also be applied in our case. In the tin structure, each atom has four nearest neighbours (at a distance of 3.016 Å) situated at the corners of a flat tetrahedron and two next nearest neighbours (3.176 Å) in the  $c$ -direction (see Fig. 4). A consideration of this structure shows that a diffusion movement of an atom in the  $c$ -direction involves two neighbouring atoms being favourably displaced from their equilibrium positions while diffusion in the  $a$ -direction requires displacement of four neighbouring atoms. This reasoning suggests that the  $A$  factor in the  $c$ -direction would be greater than that in the  $a$ -direction, since the  $A$  factor is usually regarded as a configurational term. However, the same reasoning applied to the activation energy would lead to the lower  $Q$  value in the  $c$ -direction contrary to the experimental results.

Since, unlike zinc, the tin structure is not a plane layer type it is artificial to speak of the activation energies in the two axial directions and the argument of Huntington(21) for the activation energy is inapplicable for tin.



It is not surprising that the anisotropy of such a complicated atomic process as diffusion cannot be explained simply in terms of lattice geometry. For the same reason the application of the Langmuir-Dushman equation(22) with the jump distance as main variable is not successful. However, one might expect a correlation between the activation energy and such other physical properties as are mainly determined by the energy with which the atom is bound to its equilibrium position, e.g. the linear compressibility. In tin, zinc, and bismuth crystals, this is larger in the direction of the principal axis than at right angles to it, and the same holds for the values of the coefficient of thermal expansion which depends on the ease with which the centre of the non-harmonic vibrations can be shifted with increasing amplitude. Whereas, for bismuth and zinc the direction of greater compressibility and thermal expansion coincides with that of lower activation energy,  $Q$ , this is not the case for tin and thus there seems to be no correlation between these properties.

An explanation of the anisotropy of the diffusion rate would probably be found if more was known of the mechanism of the process. The vacancy mechanism of diffusion has not been finally established, but it seems most likely that diffusion in metals occurs in this way(20). On this theory, the diffusion coefficient for a particular jump from one lattice site to a neighbouring vacant site is given(23) by

$$D = \delta^2 \nu C e^{-(J+L)/kT}, \quad \dots\dots\dots (5)$$

where  $\delta$  is the jump distance,

$\nu$  is the frequency of atomic vibrations,

$C$  is a temperature independent constant of magnitude approximately unity,

$J$  is the energy required to create a vacant site, and

$L$  is the energy barrier through which an atom must pass to reach the vacant site.

It is seen that the activation energy for a jump is made up of two terms of which  $J$  will be constant for all jumps. Variation in activation energy for different jumps will therefore arise from changes in  $L$ .

The jumps which seem most likely to occur in tin are those of an atom to a vacancy at a nearest or next nearest neighbouring site. The  $L$  values for these two jumps then determine the magnitude of the activation energies for the respective jumps. However, it is not possible at present to calculate these terms or even to predict their magnitude.

#### (b) Vacancy Concentration near the Surface

It has been remarked earlier in this paper that the first point of the diffusion-penetration curve is always high, an effect which has also been observed by other workers(9). It seems possible to explain this from the point of view of the vacancy theory of diffusion. It has been shown(23) that on the basis of this theory the diffusion coefficient for the vacancies,  $D_v$ , is given by

$$D_v = \frac{N}{n} D, \quad \dots\dots\dots (6)$$

where  $N$  is the number of atoms in the crystal,

$n$  is the number of holes in the crystal, and

$D$  is the atomic diffusion coefficient.

$n/N$  reaches a maximum of approximately  $10^{-3}$  near the melting point(24). Consequently the holes move with a velocity several orders higher than the atoms.

When any specimen is maintained at a certain temperature the equilibrium number of holes is

$$n = N C e^{-J/kT}, \quad \dots\dots\dots (7)$$

and it is assumed that this vacancy concentration is attained in a short period. This is probably so, at least near the surface, since the vacancy migration is rapid. In the centre of larger crystals equilibrium may be attained more slowly since holes must either migrate in from the surface to make up a deficiency or out to the surface to remove an excess of holes. The latter is more likely, since holes are probably frozen in on casting.\*

The occurrence of the high point in the diffusion curves can now be explained in terms of the vacancy movement of the atoms. The interface between the plated layer and the specimen will be surrounded by a region containing more vacancies than the number given by equation (7). This is for two reasons. Firstly, the original surface of the specimen will have a higher equilibrium density of holes than the bulk, due to small cracks which are inevitably present at any prepared surface, and also to the same reasons as those advanced by Anderson(25) for ionic solids. Secondly, the plated layer will have many more vacancies than the bulk of the specimen. The result of this higher vacancy density is that there is a great increase in the diffusion rate in the neighbourhood of the interface. This effect will only persist until vacancy equilibrium is established by migration and by the vacancies being eliminated by sintering together(26). In this way the concentration of radio-active atoms at short distances from the interface will be higher than that which arises from diffusion with the equilibrium number of vacancies.

## V. ACKNOWLEDGMENT

The author wishes to thank Dr. W. Boas for suggesting the problem, for his continued encouragement, and many helpful discussions throughout the work.

## VI. REFERENCES

- (1) v. HEVESY, G., SEITH, W., and KEIL, A.—*Z. Phys.* **79** : 197 (1932).
- (2) SEITH, W., and KEIL, A.—*Z. Metallk.* **25** : 104 (1933).
- (3) SEITH, W.—*Z. Elektrochem.* **39** : 538 (1933).
- (4) STEIGMAN, J., SHOCKLEY, W., and NIX, F. C.—*Phys. Rev.* **56** : 13 (1939).
- (5) MAIER, M. S., and NELSON, H. R.—*Trans. Amer. Inst. Min. (Metall.) Engrs.* **147** : 39 (1942).
- (6) MCKAY, H. A. C.—*Trans. Faraday Soc.* **34** : 845 (1938).
- (7) JOHNSON, W. A.—*Trans. Amer. Inst. Min. (Metall.) Engrs.* **143** : 107 (1941).

\* In a polycrystalline specimen the grain boundaries can also act as sources and sinks of holes and equilibrium will be rapidly established throughout.

- (8) BANKS, F. R.—*Phys. Rev.* **59**: 376 (1941).
- (9) MILLER, P. H., and BANKS, F. R.—*Phys. Rev.* **61**: 648 (1942).
- (10) BOAS, W., and FENSHAM, P. J.—*Nature* **164**: 1127 (1949).
- (11) HASLER, M. F.—*Rev. Sci. Instrum.* **4**: 656 (1933).
- (12) EVANS, U. R., and WHITWHAM, D.—*J. Electropl. Depos. Tech. Soc.* **22**: 24 (1947).
- (13) BARRETT, C. S.—“Structure of Metals.” 1st Ed. p. 175. (McGraw-Hill: New York, 1943.)
- (14) BARNES, S. W.—*Phys. Rev.* **56**: 414 (1939).
- (15) BARRER, R. M.—“Diffusion in and through Solids.” 1st Ed. Ch. I. (Cambridge University Press, 1941.)
- (16) BRIDGMAN, P. W.—*Proc. Amer. Acad. Arts Sci.* **60**: 305 (1925).
- (17) UVEN, M. J. VAN—“Mathematical Treatment of the Results of Agricultural and other Experiments.” 1st Ed. Ch. 9. (D. Noordhoff: Groningen, Batavia, 1935.)
- (18) WARING, C. E., and BECKER, P.—*J. Chem. Phys.* **15**: 488 (1947).
- (19) HINSHELWOOD, C. N.—“Kinetics of Chemical Change.” p. 257. (Oxford University Press, 1940.)
- (20) HUNTINGTON, H. B.—*Phys. Rev.* **61**: 336 (1942).
- (21) HUNTINGTON, H. B.—*Phys. Rev.* **63**: 383 (1943).
- (22) DUSHMAN, S., and LANGMUIR, I.—*Phys. Rev.* **20**: 113 (1922).
- (23) JOHNSON, R. P.—*Phys. Rev.* **56**: 814 (1939).
- (24) SEITZ, F.—“The Modern Theory of Solids.” 1st Ed. p. 458. (McGraw-Hill: New York, 1940.)
- (25) ANDERSON, J. S.—*Discuss. Faraday Soc.* **4**: 163 (1948).
- (26) SEITZ, F.—*Phys. Rev.* **74**: 1513 (1948).
- (27) JAEGER, J. C., and CARSLAW, H. S.—“Conduction of Heat in Solids.” 1st Ed. p. 30. (Oxford University Press, 1947.)

## APPENDIX I

### *Derivation of Anisotropy Relation*

Fick's law is assumed to hold for the diffusing material, i.e. the amount of material,  $F$ , diffusing across unit area in unit time is proportional to the concentration gradient.

For an anisotropic medium this can be written

$$\mathbf{F} = -\mathbf{D} \cdot \text{grad } c, \dots\dots\dots (8)$$

where  $\mathbf{D}$  is a symmetrical dyadic or second order tensor, dependent on the temperature, and  $c$  is the concentration.

Equation (8) expanded becomes

$$\left. \begin{aligned} F_x &= -\left(D_{11}\frac{\partial c}{\partial x} + D_{21}\frac{\partial c}{\partial y} + D_{31}\frac{\partial c}{\partial z}\right), \\ F_y &= -\left(D_{12}\frac{\partial c}{\partial x} + D_{22}\frac{\partial c}{\partial y} + D_{23}\frac{\partial c}{\partial z}\right), \\ F_z &= -\left(D_{13}\frac{\partial c}{\partial x} + D_{23}\frac{\partial c}{\partial y} + D_{33}\frac{\partial c}{\partial z}\right), \end{aligned} \right\} \dots\dots\dots (9)$$

where  $F_x, F_y, F_z$  are the components of the diffusion parallel to axes  $x, y, z$ .

Further the continuity principle holds for the diffusing material; and therefore

$$\text{div } \mathbf{F} = -\frac{\partial c}{\partial t} \dots\dots\dots (10)$$

Equations (10) and (8) give

$$\frac{\partial c}{\partial t} = \text{div} (\mathbf{D} \cdot \text{grad } c), \dots\dots\dots (11)$$

which is the general equation for diffusion in an anisotropic medium. For isotropic bodies (11) reduces to

$$\frac{\partial c}{\partial t} = D \nabla^2 c, \dots\dots\dots (12)$$

and for linear diffusion this may be written as

$$\frac{\partial c}{\partial t} = D \frac{\partial^2 c}{\partial x^2} \dots\dots\dots (13)$$

The  $D_r$ , in (9) are the nine components of  $\mathbf{D}$  in (8). If the axes  $Ox, Oy, Oz$  with associated unit vectors  $\mathbf{i}, \mathbf{j}, \mathbf{k}$ , are chosen parallel to the crystallographic axes in hexagonal and tetragonal lattices (e.g. Zn and Sn), the symmetry properties of these structures make all non-diagonal components of  $\mathbf{D}$  zero,

that is

$$\mathbf{D} = D_{11}\mathbf{ii} + D_{11}\mathbf{jj} + D_{33}\mathbf{kk} \dots\dots\dots (14)$$

The three non-zero components are called the principal diffusion coefficients and they will be the values of  $D$  in equation (13) when the diffusion is measured along the respective axes.

In general, the diffusion will be measured in a crystal in a direction which has direction cosines ( $l, m, n$ ) relative to the principal axes of diffusion.

The diffusion coefficient  $D_\theta$  in this direction will be given by  $\mathbf{n} \cdot \mathbf{D} \cdot \mathbf{n}$ , where  $\mathbf{n}$  is the unit vector in the diffusion direction,

$$\begin{aligned} D_\theta &= \mathbf{n} \cdot \mathbf{D} \cdot \mathbf{n} \\ &= (l\mathbf{i} + m\mathbf{j} + n\mathbf{k}) \cdot (D_{11}\mathbf{ii} + D_{11}\mathbf{jj} + D_{33}\mathbf{kk}) \cdot (l\mathbf{i} + m\mathbf{j} + n\mathbf{k}) \\ &= (lD_{11}\mathbf{i} + mD_{11}\mathbf{j} + nD_{33}\mathbf{k}) \cdot (l\mathbf{i} + m\mathbf{j} + n\mathbf{k}) \\ &= l^2D_{11} + m^2D_{11} + n^2D_{33} \\ &= D_{11}(l^2 + m^2) + D_{33}n^2 \dots\dots\dots (15) \end{aligned}$$

Since

$$l^2 + m^2 + n^2 = 1.$$

Therefore

$$l^2 + m^2 = 1 - n^2.$$

If  $n = \cos \theta$  then  $\theta$  will be the angle between the diffusion direction and the tetragonal or hexagonal axis and then

$$\begin{aligned} D_\theta &= D_{11} \sin^2 \theta + D_{33} \cos^2 \theta, \\ \text{or} \quad D_\theta &= D_a \sin^2 \theta + D_c \cos^2 \theta, \dots\dots\dots (16) \end{aligned}$$

where  $D_a, D_c$  are the diffusion coefficients normal and parallel to the tetragonal or hexagonal axis respectively.

Equation (16) has the same form as the anisotropy relation for thermal conductivity (27) which is to be expected since the differential equations governing this process are the same as equation (8).



## APPENDIX II

*Solutions used in the Preparation of Tin Crystals*

TABLE 5

## SOLUTION USED FOR ELECTROPOLISHING

Solution	Operating Conditions
144 cc. $C_2H_5OH$ 16 cc. <i>n</i> -butyl alcohol 32 cc. water 45 g. anhydrous $ZnCl_2$ 10 g. anhydrous $AlCl_3$	25 V.; 0.6 A. $cm.^{-2}$ ; 20 °C. Polish for 30 seconds, remove from solution and wash film off surface. Repeat

TABLE 6

## SOLUTION USED FOR DEVELOPING CRYSTALLOGRAPHIC ETCH PITS

Solution	Operating Conditions
5N HCl sat. with $FeCl_3 \cdot 6H_2O$ dil. with 1 part of water	Etch for 10 minutes, wiping surface every 2 minutes

TABLE 7

## SOLUTION USED FOR ELECTROPLATING

Solution	Operating Conditions
0.2 g. $Sn^{++++}$ 10 ml. conc. HCl 1 g. $NH_4Cl$ 4 g. $NH_2NH_2$ , $H_2SO_4$ Gelatin, 140 ml. water	35 °C. 1 mA. $cm.^{-2}$ for 1 hour, followed by 5 mA. $cm.^{-2}$ for 5 hours

# SELF-DIFFUSION IN POLYCRYSTALLINE TIN

By P. J. FENSHAM\*

[*Manuscript received December 14, 1949*]

## *Summary*

The rate of volume diffusion in polycrystals of tin, as indicated by the darkening of photographic films, is in general agreement with the previous measurements of the rate of self-diffusion in different crystallographic directions. The rate of diffusion along the grain boundaries is generally considerably higher than that of volume diffusion except when the boundary is that between twins or between crystals of similar orientations. These observations are discussed in terms of recent theories of the grain boundary and the vacancy mechanism of diffusion.

## I. INTRODUCTION

In a previous paper(1) the anisotropy of the rate of self-diffusion in single crystals of tin was discussed. The rate was found to depend markedly on the direction in which it was measured. For a polycrystalline specimen the diffusion rate should be some average value of the rates in the single crystals, if the rate of diffusion along the boundaries is not faster than through the bulk. The rate of boundary diffusion has been the subject of conflicting reports in the literature. Langmuir(2) found that at 2400 °K. thorium diffuses in tungsten one hundred times faster at the grain boundary than within the grains. No such difference was observed in the case of self-diffusion in lead(3) and little or none in the diffusion of nitrogen and carbon into iron(4). However, in the former case it seems likely that the effect could have been missed because of the large grain size in the polycrystals used, while the latter cases are examples of interstitial diffusion and, because of their different mechanism, are not comparable to self-diffusion in metals.

It was therefore important to see at least qualitatively whether a grain boundary effect occurred in self-diffusion of tin. The diffusion was followed by means of radioactive atoms and their effect on a photographic plate.

## II. EXPERIMENTAL METHOD AND RESULTS

Several polycrystals were grown from spectrographically pure tin (>99.998 per cent.) using the method described in the previous paper(1). From these, specimens with a flat interface for diffusion were prepared using the polishing method described before.

The following specimens were prepared :

1. A polycrystal with six large crystals in the diffusion interface.
2. A bicrystal in which the orientations of the two crystals differed widely.

\* Chemistry Department, University of Melbourne, and Division of Tribophysics, C.S.I.R.O.

3. A bicrystal in which the orientation difference was only one or two degrees.
4. A polycrystal with several twins in one of its grains.

The orientations of the crystals in specimens 2, 3, and 4 were measured by the method of etch pits(1).

A thin layer (0.02 to 0.03 mm.) of radioactive  $\text{Sn}^{113}$  mixed with inactive tin was plated on to the prepared interface and allowed to diffuse for 24 hours at 223 °C. in a vacuum ( $10^{-3}$  mm. Hg). After diffusion a thin layer normal to the diffusion direction containing the plated layer and 0.05 mm. of the original specimen was removed with a microtome. The specimen was then placed with its cut face on X-ray film for a period of 24 hours. The radiograph obtained in this way showed the distribution of radioactive atoms at this depth after diffusion.

In all cases very heavy blackening of the film around the edges of the specimen occurred, indicating extensive surface diffusion. The sides of the specimen parallel to the diffusion direction were then removed by polishing to

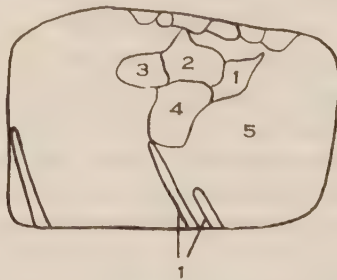


Fig. 1.—Tracing of photomicrograph of Plate 1, Figure 4b. Orientations of the crystals:  $\theta_1 = 60^\circ$ ;  $\theta_2 = 50^\circ$ ;  $\theta_3 = 53^\circ$ ;  $\theta_4 = 68^\circ$ ;  $\theta_5 = 40^\circ$ .

a depth of 0.5 mm. and a second radiograph taken. These radiographs showed much less surface darkening and are reproduced in Plate 1, Figures 1a to 4a. The cut face was then etched with acid ferric chloride solution to reveal the macro-structure, and photographs of the etched specimens are shown in Plate 1, Figures 1b to 4b.

It can be seen that different crystals blacken the film to different degrees according to the amount of radioactive material which has diffused into them, and that some of the grain boundaries appear as black lines. These general features are well shown in Plate 1, Figure 1.

In Plate 1, Figure 2, the angle,  $\theta$ , between the tetragonal axis and the diffusion direction is  $45^\circ$  in the island crystal and  $66^\circ$  in the surrounding crystal. Therefore, according to the measurements in (1) the rate of diffusion should be faster in the centre crystal, and the greater darkening of this crystal (see Plate 1, Fig. 2a) indicates that this is so.

In Plate 1, Figure 3, the interface is darkened uniformly and no grain boundary is visible. As the crystals have angles,  $\theta$ , of  $65^\circ$  and  $66^\circ$  respectively, no anisotropy effect could be expected, and the absence of a higher diffusion rate at the boundary may also be a consequence of this similarity in orientation.

In Plate 1, Figure 4, differential darkening of the grains and some grain boundaries are apparent. The twin (arrow) is seen as a thin light streak without darkening at its boundary. The orientations of the crystals are given in the legend of Figure 1 which is a tracing of Plate 1, Figure 4b. Crystal 5 is obviously the one in which the diffusion rate is highest and this is again in agreement with expectation. The small crystals on the upper edge are the result of recrystallization. No boundary is visible between crystals 2 and 3 which have almost identical  $\theta$  values. The boundary between crystals 1 and 5 is not darkened either although the rates of volume diffusion in these two crystals are obviously different. However, it was found that these crystals are within two degrees of their twin orientations, assuming with Chalmers(5) that (301) is the twin plane. It should be noted that the boundaries between these crystals are not straight and could consist of twin planes only if a stepwise nature were assumed.

### III. DISCUSSION

The experimental evidence may be summarized as follows:

- (i) The rate of volume diffusion as indicated by the darkening of the photographic films is in general agreement with the previous measurements of the rate of self-diffusion in different crystallographic directions.
- (ii) The rate of diffusion along the grain boundaries is generally considerably higher than that of volume diffusion except when the boundary is that between twins or between crystals of similar orientations.

As the grain boundary is an internal surface the atoms forming the boundary will have a higher energy than those in the interior of the crystals. The magnitude of this energy will depend on the extent of misfit of the lattices.

The relation between the grain boundary energy and orientation difference has been discussed by Dunn and Lionetti(6) and Chalmers(7). The essential feature of their results is that the energy is small at very small orientation differences and also at the twin orientation relation. For all other orientations they conclude that the boundary energy is almost independent of the relative orientations. This is also in agreement with a theoretical argument of Lennard-Jones(8).

Chalmers(9) pictures the grain boundary as an abrupt transition from the lattice of one crystal to that of the next. Kê(10) suggests that the grain boundary is rather more extended and consists of regions of order and disorder. The regions of disorder consist of small groups of atoms which form lattice imperfections of the type present in the bulk of the crystal. Between these imperfections there are regions of good lattice fit.

On both these theories there will be an equilibrium number of vacancies at an ordinary grain boundary which is several orders higher than that within the crystal. The atomic diffusion coefficient is proportional to

$$\frac{n}{N} \cdot e^{-L/kT},$$

where  $n/N$  is the ratio of vacant lattice sites and total sites, and  $e^{-L/kT}$  is the chance that an atom will have sufficient energy to move into an adjacent vacancy.



In the neighbourhood of the grain boundary both terms will be greater than in the bulk of the crystal and the diffusion rate will thus be higher.

#### IV. ACKNOWLEDGMENT

I wish to thank Dr. W. Boas for his many helpful suggestions and discussions throughout this work.

#### V. REFERENCES

- (1) FENSHAM, P. J.—*Aust. J. Sci. Res. A* **3**: 91 (1950).
- (2) LANGMUIR, I.—*J. Franklin Inst.* **217**: 543 (1924).
- (3) SEITH, W., and KEIL, A.—*Z. Metallk.* **25**: 104 (1933).
- (4) MEHL, R. F.—*Trans. A.I.M.E.I.M.D.* Lecture 122: 11 (1936).
- (5) CHALMERS, B.—*Proc. Phys. Soc.* **47**: 733 (1935).
- (6) DUNN, C. G., and LIONETTI, F.—*J. Metals* **1**: 125 (1949).
- (7) CHALMERS, B.—*Proc. Roy. Soc. A* **196**: 64 (1949).
- (8) LENNARD-JONES, J. E.—*Proc. Phys. Soc.* **52**: 38 (1940).
- (9) CHALMERS, B.—*Proc. Roy. Soc. A* **162**: 120 (1937).
- (10) KÉ, T'ING-SUL.—*Phys. Rev.* **73**: 267 (1948).

#### EXPLANATION OF PLATE 1

Fig. 1.—Specimen 1 ( $\times 4$ ).

- (a) Radiograph showing differential darkening due to different diffusion rates and outline of grain boundaries.
- (b) Photomacrograph of etched specimen.

Fig. 2.—Specimen 2 ( $\times 4$ ).

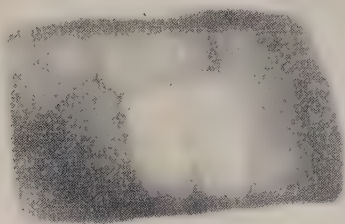
- (a) Radiograph. Differential darkening shows higher diffusion rate in the interior crystal and rapid diffusion at the grain boundary.
- (b) Photomacrograph. Island crystal:  $\theta = 45^\circ$ . Surrounding crystal:  $\theta = 66^\circ$ .

Fig. 3.—Specimen 3 ( $\times 4$ ).

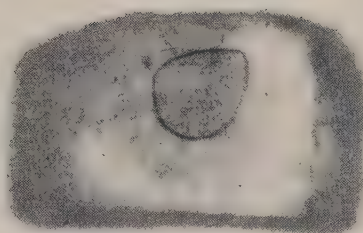
- (a) Radiograph showing uniform darkening and no outline of the grain boundary.
- (b) Photomacrograph. The crystals have almost the same orientation.

Fig. 4.—Specimen 4 ( $\times 4$ ).

- (a) Radiograph.
- (b) Photomacrograph.



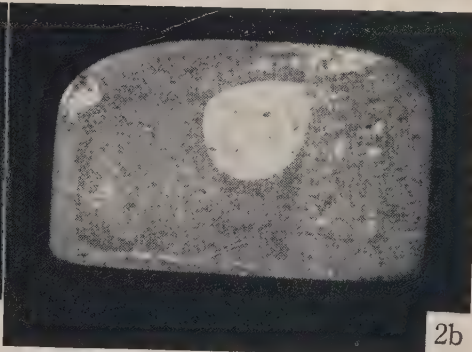
1a



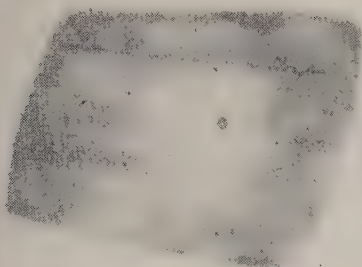
2a



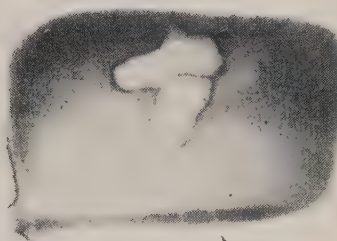
1b



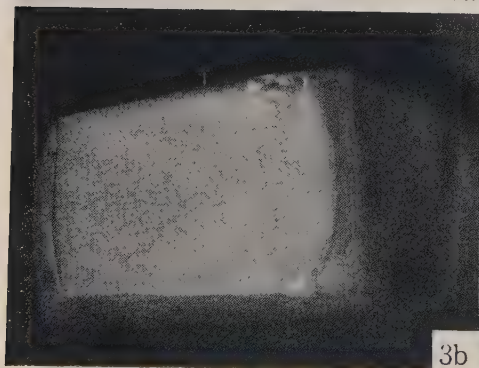
2b



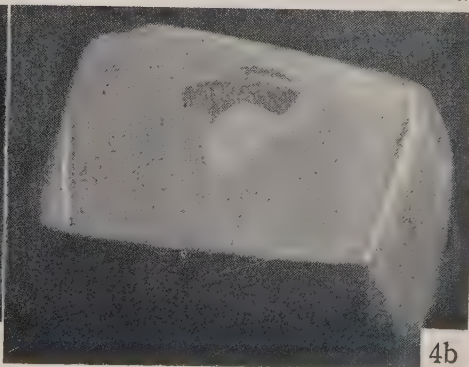
3a



4a



3b



4b



## CONJUGATED COMPOUNDS

### II. SIMPLE POTENTIAL ENERGY FUNCTIONS, ABSORPTION SPECTRA, AND IONIZATION IN LINEAR POLYENES

By N. S. BAYLISS\*

[*Manuscript received November 2, 1949*]

#### *Summary*

The  $\pi$ -electrons of a linear conjugated system are considered in terms of two simple models for the potential energy, (A) a rectangular well and (B) a "cosine-squared" well. In each case it is possible to account quantitatively for the positions of long wavelength absorption bands, their intensities, and the ionization energies in ethylene, butadiene, hexatriene, and octatetraene with the use of two parameters, the breadth and depth of the well. By comparing the models with actual periodic potential energies calculated along the long axes of polyene molecules, it is shown that an effective nuclear charge of carbon of  $Z' \approx 1$  is appropriate for the models under consideration. The assumptions involved in the application of simple one-dimensional potential energy functions to linear conjugated systems are considered in detail.

#### I. INTRODUCTION

In the first paper of this series entitled "A 'Metallic' Model for the Spectra of Conjugate Polyenes", the author(1) showed that positions and intensities in the absorption spectra of linear conjugated polyenes could be described quantitatively in terms of a metallic or free electron model, in which the unsaturation or  $\pi$ -electrons were regarded as an electron gas in a one-dimensional space coinciding approximately with the length of the molecule in question. Shortly afterwards, and independently of the author, both Kuhn(2, 3) and Simpson(4) used similar treatments for the spectra of cyanine and simple polymethine dyes. The quantum mechanics of a free electron gas in a three-dimensional space had been applied previously to the  $\pi$ -electrons of aromatic molecules in a series of papers by Schmidt(5), who, however, did not draw any quantitative conclusions about the spectra. The free electron model used by Bayliss(1) involved several simplifying assumptions of which the most important are as follows:

- (i) The  $\pi$ -electrons were considered separately from the rest of the molecule.
- (ii) The problem was confined to one dimension in the direction of the length of the polyene chain.
- (iii) The potential energy of the  $\pi$ -electrons was regarded as uniform along the conjugated chain.
- (iv) There was an infinite potential barrier at each end of the one-dimensional space (see Fig. 2 (c)).

\* Department of Chemistry, University of Western Australia, Nedlands, Western Australia.



The first assumption (i) is common to all current theoretical treatments of conjugated systems. The second (ii) means that the method in its simple form must be confined to linear molecules such as polyenes and polymethines in their all *trans*-configurations. The one-dimensional approximation is, of course, not adequate to describe aromatic compounds or linear conjugated systems in *cis*-configurations. The quantitative extension of the method to more dimensions is being investigated.\* Since the electronic potential energy along any chain of carbon atoms must be a periodic oscillating function, assumption (iii) cannot be strictly true, but has the same kind of validity as the Sommerfeld approximation that the conduction electrons in metals can be regarded as occupying a uniform potential field. Assumption (iv) greatly simplified the solution of the Schrodinger equation, but had the disadvantage that the uniform potential energy of a  $\pi$ -electron within the molecule did not have to be related to any outside energy. In other words, the model made no provision for ionization or for estimating the ionization energy.

In the present paper, while retaining assumptions (i) and (ii), it is proposed to examine the consequences of two approximate potential energy models based on the consideration of the actual form of the potential energy of a  $\pi$ -electron along the molecular axis. Apart from the fact that the potential energy must be a periodically oscillating function, it must also have the general shape of a well flattening off to a constant value of zero (corresponding to ionization) outside the molecule. The two models considered here have this property and thus permit account to be taken of the ionization energy, in contrast to the earlier model (Bayliss 1). They have been chosen as non-periodic functions to simplify the treatment. The case of actual periodic potential energy curves will be considered in a subsequent paper.†

## II. THE POTENTIAL ENERGY

We shall first consider the potential energy of a  $\pi$ -electron along the axis of a linear conjugated polyene as shown in Figure 1, where the representative molecule of butadiene is drawn to scale. At all points along the  $x$ -axis, the  $\pi$ -electron is under the electrostatic influence of the carbon and hydrogen nuclei, and the other electrons. At a point such as P, well removed from the molecule, the electron is in the field of the net residual charge  $+e$  of the rest of the molecule. If P is distant enough, the charge  $+e$  can be regarded as located at the centre of the molecule C, and the potential energy  $V_P$  is

$$V_P = -e^2/R, \dots\dots\dots (1)$$

where  $R$  is the distance PC. At a point Q within the molecule, the screening effect of the outer electrons is less, and if it is assumed

(i) that the hydrogen nuclei can be neglected,‡

\* By W. Simpson (personal communication) and by J. C. Riviere and the author.

† Part III of this series, in preparation.

‡ It is obvious from equation (2) that their effect could easily be included, with the disadvantage of introducing an additional parameter  $Z'_H$ .

(ii) that the screening effect of the other electrons can be expressed by assigning an effective nuclear charge  $Z'e$  to each carbon nucleus, then the potential energy at  $Q$  is given by

$$V_Q = \sum_k (-Z'e^2/r_k), \dots\dots\dots (2)$$

where  $r_k$  is the distance from  $Q$  to the  $k$ th carbon (see Fig. 1) and the summation

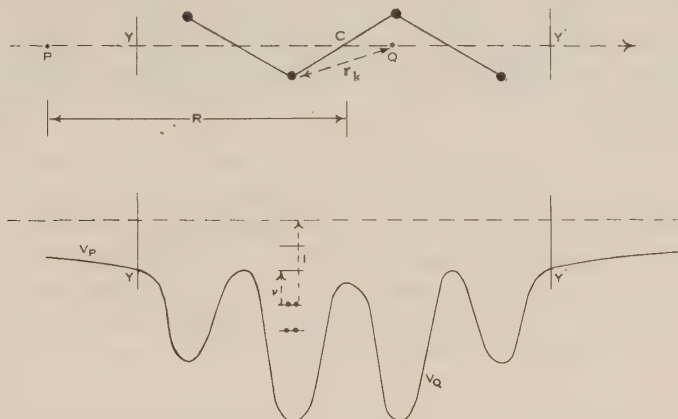


Fig. 1.—Calculation of  $\pi$ -electron potential energy along the axis of a polyene. The upper half of the figure shows the butadiene molecule to scale and the relation of the points  $YY'$  to the end carbons. The lower half of the figure shows the potential energy curve, the portions  $V_P$  outside the points  $YY'$  being calculated by equation (1), and the portion  $V_Q$  between the points  $YY'$  being calculated by equation (2).

The energy levels are diagrammatic.

is taken over all the carbons. If the bond distances and bond angles of the polyene are known, the  $r_k$  and  $V_Q$  for any point  $Q$  can be computed with little labour. As the  $\pi$ -electron under consideration passes from the representative point  $Q$  to the representative point  $P$ , the screening alters, and the form of  $V$  changes from that of equation (2) to that of equation (1). This situation can be described by joining the curves  $V_Q$  and  $V_P$  smoothly at points  $YY'$  (see Fig. 1) that lie somewhere beyond the end carbon nuclei. The final shape of the whole  $V$  curve depends on the choice of the points  $YY'$  and of the effective atomic number  $Z'$ .

The Schrodinger equation can be solved in principle for this potential energy to derive the energy levels shown diagrammatically in Figure 1. If the polyene molecule has  $2N$   $\pi$ -electrons ( $N$  conventional double bonds), the  $N$  lowest levels are occupied in the ground state, and the transitions corresponding to the main absorption band  $\nu$  and to ionization  $I$  can be represented as occurring from the highest occupied level. Experimental  $\nu$  and  $I$  data can thus be used to check the reasonableness of a potential curve of this type or of an approximation to it. A later paper will describe the numerical solution of the Schrodinger equation for actual curves of the type shown in Figure 1 by a simple form of the relaxation method of Southwell(6).

## III. TWO SIMPLE MODELS

We shall deal here with two approximations to the potential energy function  $V$  that enable the Schrodinger equation to be solved analytically. They may be described as the *rectangular potential well* (model A, Fig. 2 (a)) and the *cosine potential well* (model B, Fig. 2 (b)) respectively. They are specified more precisely as follows:

**Model A:** Within a space of length  $X$ , the  $\pi$ -electrons are in a field of uniform potential energy  $-V_0$ ; outside this space their potential energy is zero. As shown in Appendix I, this model can be solved by the "potential staircase" method.

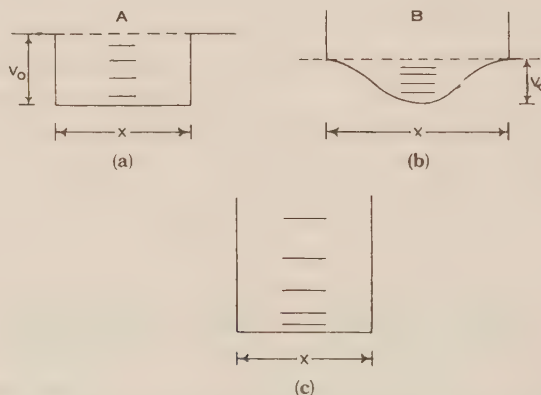


Fig. 2.—Three potential energy models for  $\pi$ -electrons in polyenes. (a) and (b) are the forms chosen for models A and B respectively, and show the significance of  $X$  and  $V_0$ . (c) shows for comparison the model used previously by the author(1).

**Model B:** Within a space of length  $X$ , the potential energy  $V$  is given by

$$V = 2k \cos^2 (\pi x/X) - V_0, \dots\dots\dots (3)$$

the depth of the well being  $V_0 = 2k$ . If the condition is imposed on this model that the wave functions shall be zero at  $x=0$  and  $x=X$ , the Schrodinger equation can be solved in terms of Mathieu's equation as shown in Appendix II.

The condition applied to model B is actually the same as that applied to the free electron model of Bayliss(1) (see Fig. 2 (c)), and in that case corresponded to an infinite ionization energy. In the present case, however, we wish to assume that an electron with  $V=0$  in equation (3) has become ionized, and that the above condition permits the problem to be "chopped off" when  $V$  has become zero at a suitable distance from the centre of the molecule. The distance  $X$  must, therefore, be rather greater than the distance  $X_c$  between the end carbon nuclei. The usual boundary condition to the Schrodinger equation is that  $\psi=0$  at  $\pm\infty$ , but in practical cases there is no objection to moving the boundary as close to the centre of the molecule as is consistent with a sensibly zero value of  $\psi$ . The relaxation method of numerically solving the Schrodinger



equation requires the problem to be chopped off at a finite distance from the centre, and for butadiene Riviere (unpublished data) has found by successive trials that the calculated energy levels are practically independent of the chopping-off point provided the potential energy and  $\psi$  have become small. In a parallel atomic case, De Groot and Ten Seldam(7) have shown that the energy levels of a "compressed" hydrogen atom are affected but little until the boundary of the problem approaches fairly closely to atomic dimensions.

In Appendices I and II, the Schrodinger equation for the respective models is solved in detail to give the energy levels and the eigenfunctions. The transition probabilities are also examined, and it is found that each model obeys the  $s \leftrightarrow a$  selection rule, namely that allowed transitions occur between a symmetrical and an antisymmetrical level, or, in other words,  $\Delta n$  must be odd where  $n$  is an integral quantum number. Formulae for the intensities of allowed transitions are developed there. In both models,  $X$  and  $V_0$  occur as parameters, and it is shown that their values can be determined uniquely from the experimental data relating to  $\nu$ , the wave number of the long wavelength absorption band, and  $I$ , the ionization energy.\* If the  $\nu$  and  $I$  data are used to fix  $X$  and  $V_0$ , the models can then be compared with experiment as follows :

- (i) The calculated length,  $X$ , should bear a reasonable relationship to known molecular magnitudes ; e.g. in model B it should be rather greater than the known distance,  $X_C$ , between the end carbons of the polyene chain.
- (ii) The calculated intensity (expressed as the oscillator strength,  $f$ ) of the main absorption band can be compared with experimental values. This calculation involves no more assumptions once  $X$  and  $V_0$  are fixed (see Appendices).
- (iii) A general matching of the calculated  $V$  curves with curves of the type shown in Figure 1 should permit a rough estimate to be made of the value of  $Z'$  in equation (2).

#### IV. COMPARISON WITH EXPERIMENTAL DATA

Experimental ionization potentials and absorption frequencies have been summarized conveniently by Price and Walsh(8) for ethylene, butadiene, hexatriene, and octatetraene, the value of  $I$  for octatetraene having been obtained by a safe extrapolation (see Table 1). By the methods outlined in Appendices I and II, these  $\nu$  and  $I$  values have been used to calculate  $X$ ,  $V_0$ , the energy levels, and the intensities,  $f$ , for each of the four molecules in terms of both models. The results of the calculations are set out in Table 2 (model A) and Table 3 (model B). It is to be noted that  $X$  agrees with molecular dimensions, and that the calculated intensities are in good agreement with the observed values. From the energy levels of Tables 2 and 3 one can recalculate values of  $\nu$  (calc.)

\* The values of  $\nu$  and  $I$  concerned are, of course, the spectroscopic "vertical" values in the Franck-Condon sense, namely those of transitions that occur without any sensible change in the linear dimensions of the molecule.



TABLE 1  
EXPERIMENTAL DATA<sup>a</sup> ON LINEAR POLYENES

	Ethylene	Butadiene	Hexatriene	Octatetraene
$\nu$ (cm. <sup>-1</sup> ) .. ..	61400	47700	39570	35000 <sup>b</sup>
$\nu$ (volts) .. ..	7.61	5.91	4.93	4.34
$I$ (volts) .. ..	10.45	9.02	8.23	7.75 <sup>c</sup>
$I/\nu$ .. ..	1.37	1.53	1.67	1.79

<sup>a</sup> Price and Walsh(8).

<sup>b</sup> Extrapolated from Price and Walsh(8). Maccoll(10) quotes 33,000 cm.<sup>-1</sup>; Woods and Schwartzmann(11) find 34,500 cm.<sup>-1</sup> in solution.

<sup>c</sup> Extrapolated value, Price and Walsh(8).

TABLE 2  
COMPARISON OF MODEL A WITH EXPERIMENT

	Ethylene $N=1$	Butadiene $N=2$	Hexatriene $N=3$	Octatetraene $N=4$
$V_0$ (volts) <sup>a</sup> ..	13.398	14.263	15.124	16.083
$X$ (calc.) <sup>a</sup> (Å) ..	2.438	4.234	5.900	7.426
$X_c$ (exp.) <sup>b</sup> (Å) ..	1.33	3.66	6.14	8.62
$f$ (calc.) .. ..	0.31	0.65	0.88	1.24
$f$ (exp.) .. ..	0.34 <sup>c</sup>	0.53 <sup>d</sup>	(0.62) <sup>e</sup>	(1.55) <sup>f</sup>
Energy levels (volts) <sup>g</sup>				
$E_1 + V_0$ .. ..	3.049	(1.34) <sup>h</sup>	(0.78)	(0.53)
$E_2 + V_0$ .. ..	10.658	5.261	(3.12)	(2.09)
$E_3 + V_0$ .. ..		11.175	6.905	(4.71)
$E_4 + V_0$ .. ..			11.835	8.295
$E_5 + V_0$ .. ..				12.635
$\nu$ (calc.) (volts) ..	7.61	5.91	4.93	4.34
$I$ (calc.) (volts) ..	10.35	9.00	8.22	7.79

<sup>a</sup> Calculated to fit experimental data.

<sup>b</sup> Experimental distance between end carbon atoms, see Bayliss(1).

<sup>c</sup> Platt, Kleven, and Price(12).

<sup>d</sup> Quoted by Mulliken and Rieke(13).

<sup>e</sup> Bayliss(1).

<sup>f</sup> Mulliken and Rieke's semi-theoretical value(13).

<sup>g</sup> Energies as measured from bottom of potential well.

<sup>h</sup> Values in parentheses obtained graphically.

and  $I$  (calc.), which should be identical with the experimental values in Table 1 if the numerical procedures have been performed accurately. The agreement is good in all cases except that of butadiene in Table 3, where the difference is

TABLE 3  
COMPARISON OF MODEL B WITH EXPERIMENT

	Ethylene $N=1$	Butadiene $N=2$	Hexatriene $N=3$	Octatetraene $N=4$
$q^a$ .. ..	2.85	6.35	12.15	19.15
$V_0$ (volts) <sup>b</sup> ..	14.702	18.218	22.862	26.442
$X$ (calc.) <sup>b</sup> (Å) ..	5.399	7.239	8.939	10.45
$X_c$ (exp.) <sup>c</sup> (Å) ..	1.33	3.66	6.14	8.62
$f$ (calc.) ..	0.32	0.74	d	d
$f$ (exp.) <sup>e</sup> ..	0.34	0.53		
Energy levels (volts) <sup>f</sup>				
$E_1$ .. ..	4.036	3.446	3.157	2.958
$E_2$ .. ..	11.666	9.873	9.205	8.616
$E_3$ .. ..		15.603	14.686	13.869
$E_4$ .. ..			19.611	18.667
$E_5$ .. ..				23.011
$\nu$ (calc.) (volts) ..	7.63	5.73	4.93	4.34
$I$ (calc.) (volts) ..	10.66	8.35	8.18	7.78

<sup>a</sup> Parameter of Mathieu equation.

<sup>b</sup> Calculated to fit experimental data.

<sup>c</sup> Experimental distance between end carbon atoms(1).

<sup>d</sup> Not calculated owing to lack of appropriate tables of Mathieu functions.

<sup>e</sup> See Table 2.

<sup>f</sup> Energies measured from lowest point of potential energy curve.

the result of errors in interpolating numerical eigenvalues of the Mathieu equation for non-integral values of its parameter  $q$ . The available tables of Mathieu functions, the best of which seem to be those of Ince(9), refer only to integral values of  $q$ . The calculated values of  $q$  for each molecule are also shown in Table 3. Tables of Mathieu functions, which are needed to compute the  $f$ 's, are similarly not available for the values of  $q$  found for hexatriene and octatetraene. Intensities for these molecules have, therefore, been calculated only for model A.

Taking butadiene as representative, the potential energy, the energy levels, and the eigenfunctions are shown to scale in Figures 3 and 4. In Figure 5, the potential energy curves of models A and B are shown superimposed on a curve for butadiene calculated from equations (1) and (2) using the bond distances and bond angles of Schomaker and Pauling(14), and assuming that the points YY' lie 0.6 Å outside the end carbons, and that  $Z'=1$ . The justification for this value will be given later.

It is of interest that both models provide only one energy level between the highest occupied level and ionization in the case of each molecule considered

(see Tables 2 and 3). This is consistent with the experimental fact that in these models no absorption band has been found between the strong ( $V, N$ ) transition and the region of the Rydberg transitions and ionization. Longer polyenes such as carotene(1) and some polymethines(4) have a weaker band at shorter wavelengths than the main band. It is proposed to consider at least the case of polymethines in a subsequent paper.

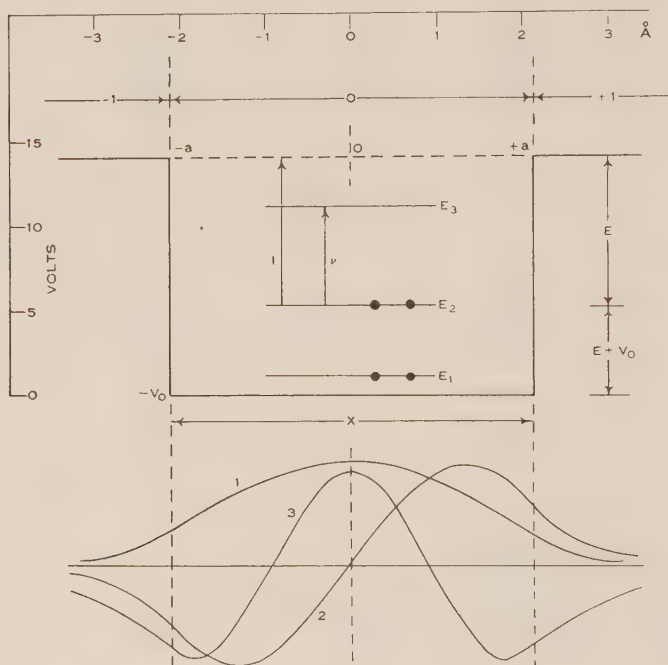


Fig. 3.—The rectangular potential energy well of model A, drawn to scale for the case of butadiene, and showing the energy levels  $E_1$ ,  $E_2$ ,  $E_3$  and the transitions  $I$  and  $\nu$ . The lower portion of the diagram shows the eigenfunctions.

## V. DISCUSSION

It is clear from the results quoted in Tables 2 and 3 that the assumption of  $\pi$ -electrons existing in a potential energy well that extends over the conjugated system is capable of giving a good quantitative account of the absorption spectra and ionization data, provided that a suitable choice is made of the parameters  $X$  and  $V_0$ . The concept of electrons in orbitals extending over a whole molecule is, of course, an essential part of all current molecular orbital theory. The present method is really a molecular orbital approach, although the latter term has been commonly applied to the LCAO approximation, from which the present treatment differs by assuming an explicit and simple form for  $V$ , the electronic potential energy. One of the greatest obstacles to a complete theoretical treatment of complex molecules is the formulation of a self-consistent expression for  $V$  in the Schrodinger equation,

$$\nabla^2\psi + (8\pi^2m/h^2)(E - V) = 0. \quad (4)$$

If this step could be achieved, the theoretical problem would be solved in principle even though the numerical difficulties would still be considerable. While significant advances towards this goal have been made recently(15),\* the most widely applied approximate methods of approach, namely the atomic orbital and the LCAO form of the molecular orbital method, avoid the explicit formulation of  $V$ . It does, however, appear in them implicitly as part of the energy operator  $H$  in the integrals of the type  $\int \psi_i^* H \psi_k d\tau$  which are so prominent in these methods. Since these integrals are nearly always evaluated by empirical

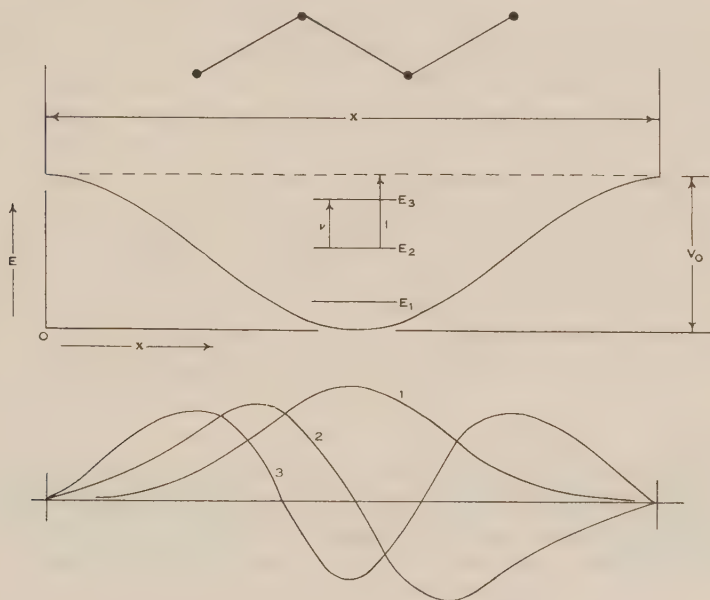


Fig. 4.—The cosine potential energy well of model B, drawn to scale for the case of butadiene, with the energy levels  $E_1$  to  $E_3$  and the transitions  $I$  and  $\nu$ . In accordance with the treatment in Appendix II, the origin of the coordinate system is shown at the lower left-hand corner. The lower portion of the figure shows the eigenfunctions.

or semi-empirical methods from experimental thermal and spectroscopic data(15–17), the question of the explicit formulation of  $V$  does not usually arise.

The present method, on the other hand, depends entirely on the choice of an admittedly approximate form for  $V$ , and once this is made the solution of the Schrodinger equation to give the energy levels and eigenfunctions follows with no further approximations. The discussion of the validity of the method can, therefore, be centred entirely on the adequacy or otherwise of the assumed form for  $V$ . It is interesting that the three simple models shown in Figures 2 (a), 2 (b), and 2 (c) all give quantitative representations of the spectroscopic

\* The author is indebted to Professor Mulliken for the opportunity of seeing this report in advance of publication.



data with reasonable values of  $X$ , and in Figures 2 (a) and 2 (b) of  $V_0$  as well, as will be shown later. Each model has a different "pattern" for its energy levels. That of Figure 2 (c)(1) has its levels increasing with  $n^2$  ( $n$  integral), that of Figure 2 (a) (model A) has energies whose spacings increase with the energy, but less rapidly than  $n^2$  (see Table 2), while in model B, Figure 2 (b), the spacings decrease as the energy increases (Table 3). In a polyene with  $2N$   $\pi$ -electrons, the fact that a model reproduces the position of the long wavelength absorption band means that it has the correct spacing between

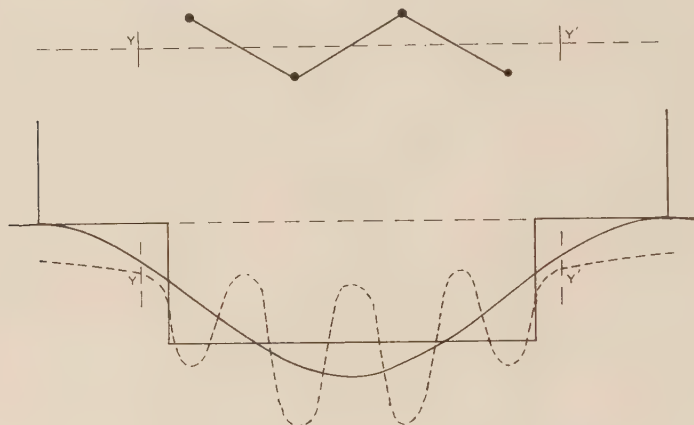


Fig. 5.—Three potential energy curves superimposed for the case of butadiene. The full curves are those of models A and B calculated to fit the experimental  $I$  and  $\nu$  data. The broken curve is the same as that of Figure 1, calculated with the assumptions that the points  $YY'$  lie  $0.6 \text{ \AA}$  outside the end carbons, and that  $Z'=1$ .

$E_N$  and  $E_{N+1}$ , where  $E_N$  is the highest occupied level in the ground state. Although the three models can be made to agree in this spacing, their other energy levels do not necessarily coincide, since all models involve different approximations to  $V$ .

As pointed out by Bayliss(1), the first model (Fig. 2 (c)) used by him and also by Kuhn(2, 3) and Simpson(4) is over-simplified in that its potential energy barrier rises vertically to infinity at each end of the system, and the resulting pattern of levels extending indefinitely upwards with  $n^2$  cannot correspond to reality, although the proper choice of the *one* parameter  $X$  makes the *one* spacing  $E_{N+1} - E_N$  agree with experiment. In view of the obvious inadequacy of this model, there does not seem to be much point in insisting, as Kuhn(2, 18) has done, on too rigid a correspondence between  $X$  and actual molecular dimensions. In treatments involving this degree of approximation, it seems better to regard  $X$  as a parameter to be determined by the experimental data. Bayliss(1) compared  $X$  with the distances  $X_c$  measured in a straight line between the ends of the polyene chain; Kuhn(3), in his treatment of the cyanines, used a length measured zig-zag fashion along the actual bonds; Simpson(4) in dealing with simple polymethines used the experimental data to determine an effective C-C bond

distance (1.28 Å) which turns out to be close to the projected distance on the axis of the average C—C distance in a zig-zag chain with  $124^\circ$  bond angles. In all of these cases the parameter  $X$  varies regularly within a homologous series of conjugated polyenes, and the addition of an extra parameter(2, 3, 4, 18) usually permits the spectral data of the whole series to be accounted for.

Kuhn(2, 18) has pointed out an important consequence of neglecting the real oscillatory character of the potential energy as shown in Figure 1. In the case of a periodic  $V$ , the energy levels are segregated into groups or Brillouin zones, the number of levels in each zone being equal to the number of periods in  $V$ . A linear polyene with  $2N$   $\pi$ -electrons has a double periodicity, a  $2N$ -fold period in the C atoms, and an  $N$ -fold period in the C=C—C units. The latter periodicity means that the  $(N, N+1)$  transition leading to the long wavelength absorption band is from the top level of the first zone containing  $N$  levels to the bottom level of the next zone, and the corresponding energy  $E_{N+1} - E_N$  is greater than if the levels were unperturbed by the periodicity of  $V$ . The value of the parameter  $X$  will therefore be different according as this  $N$ -fold periodicity is included or not in the theoretical treatment. (The effect of periodicity in  $V$  has also been treated generally by the author(19).) By considering the C=C—C periodicity, Kuhn(18) has introduced one additional parameter which accounts successfully for the way in which the first absorption maximum progresses to longer wavelengths with increasing  $N$  in the homologous series of unsymmetrical cyanines and linear polyenes.

Models A and B as described here are more realistic in that they provide for ionization, and the use of *two* parameters,  $X$  and  $V_0$ , makes it possible to fit the *two* experimental quantities  $\nu$  and  $I$ . The values of  $X$  required by model A (Table 2) are in reasonable agreement with the experimental  $X_C$ , and in the case of model B (Table 3) they are in all cases rather greater than  $X_C$ , as required by the boundary conditions. Particular interest attaches to these models, since the evaluation of the parameter  $V_0$  gives "depth" to the potential energy well and leads to the possibility of estimating the magnitude of the field in which the  $\pi$ -electrons move. The form of the periodic potential energy curve in Figure 1 depends on the choice of the effective nuclear charge  $Z'e$  of the carbon atoms and of the location of the points YY'. By roughly matching the potential energy curves of models A and B with curves of the type of Figure 1 involving different assumptions about  $Z'$  and YY', it is possible to make an estimate of  $Z'$ . The matching process was made easier by the concurrent numerical calculations on butadiene to be described in a later paper, where it will be shown that calculated and observed values of both  $\nu$  and  $I$  can be made to agree if the points YY' are about 0.6 Å outside the end carbons, and if  $Z' \approx 1$ . Figure 5 shows the three curves for butadiene superimposed in accordance with these assumptions. Visually the matching is as good as could be expected considering the different forms of potential energy concerned. It seems clear that any marked increase in  $Z'$  (say if it were doubled) would make the matching worse, since the minima of the periodic curve would be about twice as deep without much effect on the maxima. The points YY' cannot be brought much closer to the end carbons, while removing them to a greater distance would make a smaller

value of  $Z'$  necessary for good matching owing to the greater steepness of the potential energy curve  $V_Q$  (equation (2)) as compared with  $V_P$  (equation (1)). Similar comparisons in the cases of ethylene, hexatriene, and octatetraene also lead to the conclusion that  $Z' \approx 1$ .

The value of  $Z'$  indicated by the admittedly rough matching of potential energy curves in Figure 5 depends on several assumptions as well as that of choosing the location of  $YY'$ . They are:

- (i) The validity of the one-dimensional approximation to a three-dimensional molecule.
- (ii) That the hydrogens can be neglected.
- (iii) That no serious error is involved in omitting the periodic nature of  $V$  in models A and B.
- (iv) That the potential energy calculated along the molecular axis as in Figure 1 is the best representation of the potential energy of the  $\pi$ -electrons.

In the case of (ii), the periodic  $V$  curve of Figure 1 or Figure 5 is to be regarded as a curve which gives the correct  $\nu$  and  $I$ . If, then, a set of terms involving  $Z'_H$  is added to equation (2), the value required for  $Z'$  of carbon would be smaller than the one suggested. Assumption (iii) does not have any significant effect on the depth of the potential energy curve in its relation to  $\nu$  and  $I$ . This point will be made clear in a later paper, where calculations on actual periodic  $V$  curves will show that the same value of  $Z' \approx 1$  is required for butadiene.

Assumption (iv) depends on the one-dimensional approximation (i) in that we have really assumed the molecular axis in Figure 1 to be representative of the average position of the  $\pi$ -electrons. To the extent that a one-dimensional approximation is valid at all, and regarding the  $\pi$ -electrons (or their orbitals) as being located in the outer part of the electron cloud of the molecule, there will be some line parallel to the molecular axis that is the best representation of their location. This line is not necessarily the molecular axis itself; it might be a line parallel to it but displaced in a direction normal to the molecular plane, i.e. lying above or below the plane of Figure 1. If we had chosen such a line to calculate  $V_Q$  by equation (2), the  $r_k$  would have been greater, and the corresponding  $Z'$  to match the experimental data would also have been greater. The choice of the molecular axis as representative of the  $\pi$ -electrons has at least the advantage that it is a geometrically definite location. Whether it is the best choice cannot really be decided until it is possible to solve the problem for a self-consistent field in three dimensions.

Zener(20) and Slater(21) have shown that the value of  $Z'$  for a valency shell electron in an isolated carbon atom is about 3.2, and the same value has been used for  $\pi$ -electrons in conjugated compounds, together with the assumption of neglecting all but nearest-neighbour interactions, in theoretical treatments such as that of benzene by Goeppert-Mayer and Sklar(22). According to Slater(21), the value of  $Z'$  in an isolated atom is derived from  $Z$  (the atomic number) by subtracting a screening constant of 0.35 for each other electron in the same shell as the one under consideration, of 0.85 for each electron in the next inner



shell, and of 1.00 for electrons in further inner shells. If butadiene in terms of molecular orbital theory is considered as having an electron cloud of successive shells consisting of 4  $\pi$ -electrons, 12  $\sigma$ -bonded electrons, and 8  $K$  electrons, then the application of Slater's rules to the present case would give

$$4Z' = 4Z - (8 \times 1.00) - (12 \times 0.85) - (3 \times 0.35), \dots\dots\dots (5)$$

resulting in  $Z' = 1.19$ . Similar values of  $Z' \approx 1$  are also found for ethylene, hexatriene, and octatetraene.

While this result(23) is in very good agreement with the value of  $Z'$  found by matching the potential energy curves, it must be remembered that the coincidence is considerably dependent, as shown above, on the choice of the molecular axis as the best representation of the  $\pi$ -electrons. If the complete three-dimensional approach showed that some other line through the molecule was a better representation, then the appropriate value of  $Z'$  would be different. Work in progress\* shows that the application of the Slater screening constants to the carbon and nitrogen atoms in the linear polymethine ion  $(\text{NH}_2\text{.CH:CH.CH:NH}_2)_+$  leads to reasonable (although not exact) agreement between experiment and the result of the relaxation method applied to a potential energy curve of the type of Figure 1. Similar calculations on benzene show that the value of  $Z'$  depends on the choice of the location of that plane which is the best representation of the  $\pi$ -electrons.

In spite of the uncertainty about the "best" value of  $Z'$  which is implied by the preceding paragraphs, it is possible to deduce the important consequence that the "atomic" value of  $Z' = 3.2$  is not necessarily the correct value where the  $\pi$ -electrons of conjugated compounds are concerned. This result is not surprising when it is considered that the atomic value includes the assumption that the  $2s$  and  $2p$  electrons have the same screening effect(21), whereas in conjugated systems it is realized universally that the unsaturation or  $\pi$ -electrons are in a different category from the  $\sigma$ - and  $K$ -electrons.

## VI. ACKNOWLEDGMENTS

The work was facilitated by a calculating machine purchased with the assistance of the Commonwealth Research Grant to Australian Universities. The author is indebted to Mr. R. H. Leary and to Mr. J. C. Riviere for assistance with the calculations.

## VII. REFERENCES

- (1) BAYLISS, N. S.—*J. Chem. Phys.* **16** : 287 (1948).
- (2) KUHN, H.—*J. Chem. Phys.* **16** : 840 (1948).
- (3) KUHN, H.—*Helv. Chim. Acta* **31** : 1441 (1948).
- (4) SIMPSON, W. T.—*J. Chem. Phys.* **16** : 1124 (1948).
- (5) SCHMIDT, O.—Summarized in *Ber. dtsch. Chem. Ges.* **73A** : 97 (1940).
- (6) SOUTHWELL, R. V.—"Relaxation Methods in Theoretical Physics." (Oxford Univ. Press, 1946.)
- (7) DE GROOT, S. R., and TEN SELDAM, C. A.—*Physica* **12** : 669 (1946).

\* N. S. Bayliss and J. C. Riviere, in preparation.



- (8) PRICE, W. C., and WALSH, A. D.—*Proc. Roy. Soc. A* **185** : 182 (1946).
- (9) INCE, E. L.—*Proc. Roy. Soc. Edin.* **52** : 355 (1932).
- (10) MACCOLL, A.—*Quart. Rev. (Chem. Soc. Lond.)* **1** : 52 (1947).
- (11) WOODS, G. F., and SCHWARTZMANN, L. H.—*J. Amer. Chem. Soc.* **71** : 1396 (1949).
- (12) PLATT, J. R., KLEVEN, H. B., and PRICE, W. C.—*J. Chem. Phys.* **17** : 466 (1949).
- (13) MULLIKEN, R. S., and RIEKE, C. A.—*Rep. Progr. Phys.* **8** : 231 (1941).
- (14) SCHOMAKER, V., and PAULING, L.—*J. Amer. Chem. Soc.* **61** : 1769 (1939).
- (15) MULLIKEN, R. S.—Report on molecular orbital theory to be published in *J. Chim. Phys.* **46** : 497 (1949).
- (16) MULLIKEN, R. S., and RIEKE, C. A.—*Rev. Mod. Phys.* **14** : 259 (1942).
- (17) COULSON, C. A., and LONGUET-HIGGINS, H. C.—*Proc. Roy. Soc. A* **191** : 39 (1947).
- (18) KUHN, H.—*Z. Elektrochem.* **53** : 165 (1949).
- (19) BAYLISS, N. S.—*Aust. J. Sci.* **12** : 12 (1949).
- (20) ZENER, C.—*Phys. Rev.* **36** : 54 (1930).
- (21) SLATER, J. C.—*Phys. Rev.* **36** : 58 (1930).
- (22) GOEPPERT-MAYER, M., and SKLAR, A. L.—*J. Chem. Phys.* **16** : 645 (1948).
- (23) BAYLISS, N. S.—*J. Chem. Phys.* **17** : 1353 (1949).
- (24) FRENKEL, J.—“Wave Mechanics—Elementary Theory.” p. 107. (Oxford Univ. Press, 1932.)
- (25) BICKLEY, W. G.—“Mathematical Tables and other Aids to Computers.” Vol. I, p. 409 (1945).

## APPENDIX I

Figure 3 shows the significance of the symbols employed in the solution of model A. The solution of the Schroedinger equation (4) is subject to the conditions

$$\left. \begin{aligned} V &= -V_0 \text{ for } -a < x < a \quad (X=2a), \\ V &= 0 \text{ for } |x| > a, \\ E &< 0 \text{ (bound electronic levels).} \end{aligned} \right\} \dots\dots\dots (6)$$

$V_0$  is a positive quantity; the energy levels  $E$  are negative and are measured relative to the energy of a free electron at infinity. Energy levels measured from the bottom of the potential well are given by the positive quantities  $(E + V_0)$ . The three regions in Figure 3 from  $-\infty$  to  $-a$ ,  $-a$  to  $+a$ ,  $+a$  to  $+\infty$ , are indicated by the subscripts  $-1, 0, +1$  respectively.

(i) *The Energy Levels.*—Solutions of the problem in standard treatises (24) show that in the region 0

$$\psi_0 = A_0 \cos(\alpha x + \phi), \dots\dots\dots (7)$$

and in the regions  $\pm 1$

$$\psi_{\pm 1} = A_{\pm 1} e^{\mp \beta x} \dots\dots\dots (8)$$

$\psi_0$  and  $\psi_{\pm 1}$  are subject to the continuity conditions

$$\text{if } x = \pm a : \psi_0 = \psi_{\pm 1}; \quad d\psi_0/dx = d\psi_{\pm 1}/dx. \dots\dots\dots (9)$$

The parameters  $\alpha$  and  $\beta$  are positive and contain the energy levels since by substitution of equations (7) and (8) in the Schroedinger equation one obtains

$$\alpha^2 = (8\pi^2 m / h^2)(E + V_0); \quad \beta^2 = -(8\pi^2 m / h^2)E \dots\dots\dots (10)$$

From the continuity conditions of equation (9) one finds that  $\varphi=0$  or  $\pi/2$ , resulting in two sets of solutions:

<i>Symmetrical levels</i>	<i>Antisymmetrical levels</i>
$\varphi=0$	$\varphi=\pi/2$
$\psi_0=A_0 \cos \alpha x$	$\psi_0=A_0 \sin \alpha x$ ..... (11)
$A_{+1}=A_0 e^{\beta a} \cos \alpha a$	$A_{+1}=A_0 e^{\beta a} \sin \alpha a$ ..... (12)
$A_{-1}=A_{+1}$	$A_{-1}=-A_{+1}$ ..... (13)
$\beta/\alpha=\tan \alpha a$	$\beta/\alpha=-\cot \alpha a$ ..... (14)

The eigenvalues of  $\alpha$  and  $\beta$  (i.e. of  $E$ ) are found conveniently (24) by introducing the parameters  $\xi$  and  $b$  such that

$$\xi=\alpha a; \quad \xi^2=(8\pi^2 m a^2/h^2)(E+V_0) \quad \text{..... (15)}$$

$$b^2-\xi^2=\beta^2 a^2; \quad b^2=(8\pi^2 m a^2/h^2)V_0 \quad \text{..... (16)}$$

Equations (14) then become

$$\text{Symmetrical levels:} \quad \tan \xi = \sqrt{(b^2-\xi^2)/\xi^2} \quad \text{..... (17)}$$

$$\text{Antisymmetrical levels:} \quad -\cot \xi = \sqrt{(b^2-\xi^2)/\xi^2} \quad \text{..... (18)}$$

Since the value of  $b$  is determined (equation (16)) by the constants  $a$  and  $V_0$  of the model, the energy levels ( $E+V_0$ ) are found directly from the values of  $\xi$  that satisfy equations (17) and (18), which can be solved without difficulty by trial and error or graphically.

(ii) *Relation to Experimental Data.*—Denote the energy levels and the corresponding  $\xi$ 's by  $E_n$  and  $\xi_n$  where  $n=1, 2, 3, \dots$ ,  $E_1$  is the lowest level, and levels are symmetrical when  $n$  is odd, antisymmetrical when  $n$  is even. In a conjugated polyene with  $2N$   $\pi$ -electrons, levels  $E_1$  to  $E_N$  are occupied, and the wave number  $\nu$  of the first absorption band is

$$\nu \text{ (cm.}^{-1}\text{)} = (E_{N+1} - E_N)/hc = (h/8\pi^2 mca^2)(\xi_{N+1}^2 - \xi_N^2) \quad \text{..... (19)}$$

The ionization energy  $I$  (in  $\text{cm.}^{-1}$ ) from the highest occupied level  $E_N$  is given by

$$I \text{ (cm.}^{-1}\text{)} = -E_N/hc = (h/8\pi^2 mc)a\beta^2, \quad \text{..... (20)}$$

or, substituting from equation (16),

$$I = (h/8\pi^2 mca^2)(b^2 - \xi_N^2) \quad \text{..... (21)}$$

From equations (19) and (21) one obtains

$$I/\nu = (b^2 - \xi_N^2)/(\xi_{N+1}^2 - \xi_N^2) \quad \text{..... (22)}$$

There is a *unique* value of  $b$  (with its corresponding set of  $\xi$ 's determined by equations (17) and (18)) which will satisfy a given experimental value of  $I/\nu$  in terms of equation (22). This value of  $b$  may be found by trial and error, a procedure which was found to be surprisingly rapid on a calculating machine. Once  $b$  and its  $\xi$ 's are so found, equation (19) provides the value of  $a$  (and hence of  $X$ ) by using the experimental  $\nu$ .

The experimental values of  $I$  and  $\nu$  for a given polyene thus make it possible to determine uniquely the appropriate values of  $X$ ,  $V_0$ , and the energy levels of model A shown in Figure 3.

(iii) *Normalization of the Eigenfunctions*  $\psi_n$ .—It is convenient to write

$$S = \int \psi^2 dx \quad \text{..... (23)}$$

with subscripts  $S_{-1}$ ,  $S_0$ ,  $S_{+1}$  to denote the regions in Figure 3. For region 0,

$$S_0 = \int_{-a}^a A_0^2 \cos^2 \alpha x dx \text{ or } \int_{-a}^a A_0^2 \sin^2 \alpha x dx \dots\dots\dots (24)$$

for the symmetrical and antisymmetrical levels respectively.

These reduce to

$$S_0 = A_0^2 a + (A_0^2/2\alpha) \sin 2\alpha a \text{ or } A_0^2 a - (A_0^2/2\alpha) \sin 2\alpha a \dots\dots (25)$$

It is also easy to show that

$$S_{0\pm 1} = (A_{\pm 1}^2/2\beta) e^{-2\beta a} \dots\dots\dots (26)$$

The condition for normalization is that

$$S = S_{-1} + S_0 + S_{+1} = 1 \dots\dots\dots (27)$$

Using the relations between  $A_0$  and  $A_{\pm 1}$  given by equations (12) and (13), and putting  $\alpha = \beta/\tan \alpha a$  or  $-\beta/\cot \alpha a$  from equation (14), it is not difficult to show that for both symmetrical and antisymmetrical levels

$$S = A_0^2(a + 1/\beta), \dots\dots\dots (28)$$

and hence

$$A_0 = (a + 1/\beta)^{-1/2}, \dots\dots\dots (29)$$

$$A_{+1} = A_{-1} = (a + 1/\beta)^{-1/2} e^{\beta a} \cos \alpha a \text{ (symmetrical levels), } \dots\dots (30)$$

$$A_{+1} = -A_{-1} = (a + 1/\beta)^{-1/2} e^{\beta a} \sin \alpha a \text{ (antisymmetrical levels).} \dots\dots\dots (31)$$

Normalized eigenfunctions for butadiene are shown in Figure 3.

(iv) *Selection Rules and Intensities*.—To obtain the transition probability between levels with quantum numbers  $m$  and  $n$  it is required to find the transition integral

$$Q = \int_{-\infty}^{\infty} \psi_m x \psi_n dx \dots\dots\dots (32)$$

Symmetry considerations show that  $Q=0$  if levels  $m$  and  $n$  are both symmetrical or both antisymmetrical, since it is obvious from Figure 3 that in either case every term of  $Q$  where  $x > 0$  is cancelled by an equal term of opposite sign where  $x < 0$ . We thus have the well-known  $s \leftrightarrow a$  selection rule; in other words,  $|m-n|$  must be *odd* for an allowed transition.

To compute the intensities of allowed transitions, suppose that  $m$  is an *antisymmetrical*,  $n$  a *symmetrical*, level.

In the region 0,

$$\psi_m = A_{0m} \sin \alpha_m x; \quad \psi_n = A_{0n} \cos \alpha_n x, \dots\dots\dots (33)$$

$$Q_0 = A_{0m} A_{0n} \int_{-a}^a x \sin \alpha_m x \cos \alpha_n x dx \dots\dots\dots (34)$$

It may be proved that

$$Q_0 = A_{0m} A_{0n} \left[ \frac{\sin (m-n)a}{(m-n)^2} + \frac{\sin (m+n)a}{(m+n)^2} - \frac{a \cos (m-n)a}{(m-n)} - \frac{a \cos (m+n)a}{(m+n)} \right] \dots\dots\dots (35)$$

In the regions  $\pm 1$ ,

$$\psi_m = A_{\pm m} e^{\mp \beta_m x}; \quad \psi_n = A_{\pm n} e^{\mp \beta_n x}, \dots \quad (36)$$

$$Q_{+1} = A_{+m} A_{+n} \int_a^\infty x e^{-\gamma x} dx \dots \quad (37)$$

$$Q_{-1} = A_{-m} A_{-n} \int_{-\infty}^{-a} x e^{\gamma x} dx, \dots \quad (38)$$

where  $\gamma = \beta_m + \beta_n$ . Remembering that  $A_{+m} = -A_{-m}$ , and  $A_{+n} = A_{-n}$ , one finds that

$$Q_{+1} + Q_{-1} = A_{+m} A_{+n} \frac{e^{-\gamma a}}{\gamma^2} (\gamma a + 1), \dots \quad (39)$$

and introducing the relation between  $A_0$  and  $A_{\pm 1}$ ,

$$Q_{+1} + Q_{-1} = (2A_{0m}A_{0n}/\gamma^2) \sin \alpha_m a \cos \alpha_n a (\gamma a + 1) \dots \quad (40)$$

For calculation, equations (35) and (40) are combined to give the total  $Q$ , which is related to the oscillator strength  $f$  of the transition by the well-known relation

$$f = 1.08 \times 10^{11} \nu Q^2 \dots \quad (41)$$

The quantities  $A_{0m}$ ,  $A_{0n}$ ,  $\alpha_m$ ,  $\alpha_n$ ,  $\gamma = \beta_m + \beta_n$  are shown by the analysis in the previous sections to be fixed by the model parameters  $a$  and  $V_0$ .

## APPENDIX II

In discussing model B, it makes for simplicity in notation to place the zero of  $x$  at the left-hand boundary of the problem, and the zero of  $V$  at the lowest point of the potential energy curve, as shown in Figure 4. The range of  $x$  is from 0 to  $X$ , and the potential energy  $V_0$  is supposed to correspond to ionization. With this change of coordinates,  $V$  takes the form

$$V = 2k \cos^2 (\pi x/X), \dots \quad (42)$$

where  $V_0 = 2k$ . This is identical with equation (3) except for the change of zero. Putting  $\theta = \pi x/X$ , the Schrodinger equation becomes

$$d^2\psi/d\theta^2 + (8mX^2/\hbar^2)(E - k - k \cos 2\theta)\psi = 0, \dots \quad (43)$$

which is identical with the canonical form of Mathieu's equation as used by Ince(9) but in Bickley's(25) notation, namely

$$d^2y/dz^2 + (a - 2q \cos 2z)y = 0 \dots \quad (44)$$

if the following relations hold:

$$(8mX^2/\hbar^2)(E - k) = a, \dots \quad (45)$$

$$(8mX^2/\hbar^2)k = 2q. \dots \quad (46)$$

If the eigenfunctions  $\psi$  are required to be zero at the points  $x=0$  and  $x=X$ , i.e. if the problem is "chopped off" at these two points, the Mathieu equation is soluble in terms of the functions  $se_n(q, \theta)$  for eigenvalues of  $a$  which are written in conventional notation for the  $se$  functions as  $b_n$  ( $n=1, 2, 3, \dots$ ) and which are functions of the parameter  $q$ . The energy levels  $E_n$ , therefore, become

$$E_n = (\hbar^2/8mX^2)b_n + k \dots \quad (47)$$

As before, the highest occupied level in a linear polyene with  $2N$   $\pi$ -electrons is  $E_N$ , and the wave number  $\nu$  of the main absorption band is

$$\nu \text{ (cm.}^{-1}\text{)} = (E_{N+1} - E_N)/\hbar c = (b_{N+1} - b_N)(\hbar/mcX^2), \dots \quad (48)$$



while the ionization energy  $I$  in  $\text{cm}^{-1}$  is given by

$$I = V_0 - E_N / hc = (2q - b_N)(h/mcX^2) \dots\dots\dots (49)$$

From equations (48) and (49) one obtains

$$I/\nu = (2q - b_N)/(b_{N+1} - b_N) \dots\dots\dots (50)$$

Extensive tables of the  $b_n$  as functions of  $q$  are given by Ince(9), and it is found that for a given experimental value of the ratio  $I/\nu$ , equation (50) can be solved graphically to give a unique value of  $q$ , and hence of  $b_N$  and  $b_{N+1}$ . Equation (48) with the experimental value of  $\nu$  then gives  $X$ , and  $V_0 (=2k)$  is then obtained from equation (46).

The constants of model B can thus be fixed uniquely by the  $I$  and  $\nu$  data in any given case.

(i) *Normalization of the Eigenfunctions.*—The eigenfunctions of equation (43) are

$$\psi_n = C \text{se}_n(\theta) = C \text{se}_n(\pi x/X); \quad n=1, 2, 3, \dots\dots\dots (51)$$

where  $C$  is the normalizing factor to make

$$S = C^2 \int_0^X \text{se}^2(\pi x/X) dx = 1 \dots\dots\dots (52)$$

Putting  $\theta = \pi x/X$ , this means that

$$S = (C^2 X/\pi) \int_0^\pi \text{se}^2 \theta d\theta = 1 \dots\dots\dots (53)$$

These functions tabulated by Ince(9) are normalized such that

$$\int_0^\pi \text{se}^2 \theta d\theta = \pi/2, \dots\dots\dots (54)$$

from which  $C = \sqrt{(2/X)}$ , and the normalized eigenfunctions are

$$\psi_n = \sqrt{(2/X)} \text{se}_n(\pi x/X), \dots\dots\dots (55)$$

the  $\text{se}_n$  being read for the value of  $q$  appropriate to the molecule in question.

(ii) *Selection Rules and Intensities.*—For the transition probability between levels of quantum numbers  $m$  and  $n$  we again require a transition integral  $Q$  of the type  $Q = \int \psi_m x \psi_n dx$ . While so far in this problem the choice of the zero for  $x$  has been immaterial, a moment's consideration will show that the zero for the integral  $Q$  should be the centre of the molecule\* where  $x = \frac{1}{2}X$ . This

\* Providing the Schroedinger equation is soluble at all, the eigenvalues and the eigenfunctions must be independent of the choice of the origin. The origin for equation (42) was chosen to make equation (43) readily conform to Ince's tables and form of the Mathieu equation. Had we chosen the origin in the centre of the molecule, the eigenfunctions would have appeared in the form of  $\text{se}_n \theta$  ( $n$  odd; symmetrical levels) and  $\text{se}_n \theta$  ( $n$  even; antisymmetrical levels) with appropriate changes in the sign of  $q$ , but without changing the numerical values of the eigenfunctions or eigenvalues. It is readily seen, however, that the value of  $Q$  in general, although not necessarily in particular cases, depends on the origin of  $x$ .

The author(1) previously (uniform  $V$ , sine form of eigenfunctions), committed the error of computing  $Q$  with the origin at the left hand of the problem. Fortunately in that case the wrong assumption made no difference to the prohibition of transitions for which  $|m-n|$  is even, and only a negligible difference to the intensities of transitions for which  $|m-n|$  is odd, since only the sign of the minor term in the expression for  $Q$  was affected.

condition can be met in terms of our chosen coordinate system by writing

$$Q = (2/X) \int_0^X (x - X/2) \text{se}_n(\pi x/X) \text{se}_m(\pi x/X) dx \dots\dots (56)$$

Putting  $\theta = \pi x/X$ , then

$$Q = (2X/\pi^2) \int_0^\pi (\theta - \pi/2) \text{se}_n\theta \text{se}_m\theta d\theta \dots\dots\dots (57)$$

From symmetry considerations, as in Appendix I,  $Q=0$  if the levels  $m$  and  $n$  are both symmetrical or both antisymmetrical with respect to the centre of the molecule. Allowed transitions are subject to the  $s \leftrightarrow a$  selection rule, i.e.  $|m-n|$  is odd. In these cases  $Q$  can be evaluated numerically by a suitable method such as Simpson's rule. The oscillator strength  $f$  is then calculated by equation (41).

# A COLORIMETRIC METHOD FOR THE DETERMINATION OF ESTERS

By ADRIENNE R. THOMPSON\*

[Manuscript received November 11, 1949]

## Summary

The colorimetric method of Hill, based on conversion of esters to hydroxamic acids which form red complexes with ferric ion, has been adapted for the estimation of volatile esters.

Investigation of the factors influencing the two reactions involved has led to a satisfactory procedure for esters derived from acids with an unbranched chain of two to ten carbon atoms and esters of isobutyric and isovaleric acids. The degree of conversion to ferric hydroxamate complexes has been found by reference to pure hydroxamic acids.

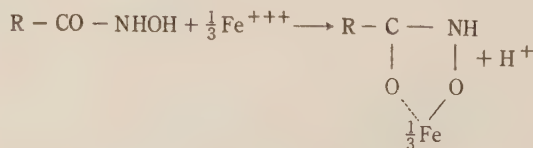
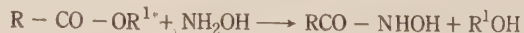
The absorption spectra of the ferric hydroxamate complexes from all the esters examined show a maximum absorption at 520 mμ. Hence a mixture of esters may be estimated in terms of one of the esters.

## I. INTRODUCTION

In the course of an investigation of the volatile substances given off to the atmosphere by apples and their relation to Superficial Scald, a physiological disorder of cold stored fruit, the need arose for a sensitive method for estimating esters. The amounts of esters which could be collected were of the order of 0.1 mg. and were far too small to allow the customary hydrolysis-titration methods(1) to be used.

A spot test described by Feigl(2) was developed into a quantitative colorimetric method for determining fatty acid esters by Hill(3, 4). The test was also used by Keenan(5) for the estimation of ethyl acetate, and by Bauer and Hirsh(6) for the determination of esterified fatty acids in human sera. The method has now been modified to provide a satisfactory procedure for the determination of volatile esters.

The basis of the determination is the reaction of the ester with hydroxylamine in alkaline solution to form a hydroxamic acid which, in turn, reacts with ferric ion to form a red ferric hydroxamate complex.



\* Division of Food Preservation and Transport, C.S.I.R.O., Homebush, N.S.W.

In the procedure developed by Hill(3, 4), which has not previously been applied to volatile esters, the first reaction was carried out in ether, which was evaporated at 70 °C. during the reaction. The second reaction was carried out in ethanol solution by the addition of ferric perchlorate.

Preliminary experiments using Hill's procedure and ethyl acetate as the ester, showed that the reaction was affected by several factors, which included the volume of the ether solution, the time of heating, the time of delay before adding ferric ions, and the temperature of the solution at the time of addition of the ferric ions. Moreover, evaporation of the ether at 70 °C. led to a loss of volatile esters.

The most satisfactory procedure for the estimation of volatile esters was found to be formation of the hydroxamic acids in ether solution at 25 °C. and conversion to the red ferric hydroxamate complex without evaporation of the ether.

## II: CONDITIONS INFLUENCING THE REACTION

Various factors which influence the rate of the reaction and the final intensity of the colour were investigated. The colour of the ferric hydroxamate complex was determined by means of an Evelyn colorimeter using a 520 filter, and expressed as density ( $2 - \log$  per cent. transmission). This figure was corrected by subtracting the density of the blank.

The conversion of esters to hydroxamic acids was found to proceed most rapidly in ether, less rapidly in dioxane, and least in ethanol (Fig. 1). Although ethanol was quite suitable for the acetic esters, it failed to promote a sufficiently rapid conversion of the esters of butyric and higher acids. The reaction in ethanol could be accelerated by increasing the concentration of hydroxylamine, but this interfered with the subsequent colour reaction.

The presence of water in the ether (up to a concentration of 1.3 per cent.) had no effect on the reaction. Ether on standing gave highly coloured blanks. This effect was probably due to the development of peroxide (Reimers 7). It could be prevented by the addition of diphenylamine (0.001–0.01 per cent.) to the ether and storage in dark bottles in a dark cupboard as suggested by Reimers. The alkaline hydroxylamine solution was prepared immediately before use by mixing methanolic solutions of hydroxylamine hydrochloride and sodium hydroxide. If the precipitated sodium chloride was not removed, it dissolved subsequently on addition of the alcoholic ferric perchlorate and reduced the final colour. The small amount of sodium chloride remaining dissolved in the hydroxylamine solution was found to be without effect on the colour.

The hydroxylamine solution was found to be stable for at least twenty minutes after preparation.

## III. RECOMMENDED PROCEDURE

### (a) Reagents

(i) *Solvents*.—Ethanol and methanol are freed from esters and aldehydes by refluxing for two days with zinc dust and sodium hydroxide and redistilling.

Commercial ether is freed from esters by refluxing over concentrated sodium hydroxide solution for three hours, separating, and distilling. A sample of



anaesthetic ether was found to be satisfactorily free from esters and peroxide without purification. The ether is stabilized against peroxide formation by the addition of 0.01 per cent. diphenylamine.

(ii) *Standard Ester Solution*.—The esters are purified by washing with sodium carbonate solution, water, and calcium chloride solution, separating, drying over calcium chloride, and fractional distillation. Stock solutions of esters ( $4 \times 10^{-2}$  M) in ethanol are prepared. Reagent ester solutions ( $8 \times 10^{-4}$  M) are made by dilution of the stock solutions with ether.

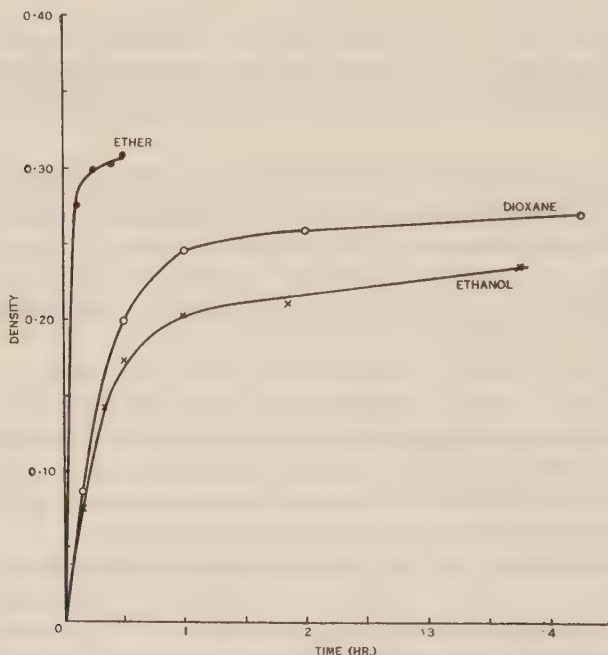


Fig. 1.—Formation of hydroxamic acid from ethyl butyrate in different solvents at 25 °C.

The reaction mixture contained 2 ml. ethyl butyrate ( $8 \times 10^{-4}$  M) in ether, dioxane, or ethanol and 0.6 ml. hydroxylamine solution (made by mixing equal volumes of 12.5 per cent.  $\text{NH}_2\text{OH} \cdot \text{HCl}$  and 12.5 per cent.  $\text{NaOH}$  in methanol).

(iii) *Hydroxylamine Solution*.—A solution of hydroxylamine containing sodium hydroxide is prepared from equal volumes of 5 per cent. hydroxylamine hydrochloride and 12.5 per cent. sodium hydroxide in methanol. The precipitated sodium chloride is removed by filtration. The solution takes about five minutes to prepare, and should be made immediately before use.

(iv) *Ferric Perchlorate Reagent Solution*.—0.4 g. iron or its equivalent as ferric chloride is dissolved in 5 ml. concentrated hydrochloric acid; 5 ml. 70 per cent. perchloric acid is added, the solution evaporated almost to dryness, and diluted to 100 ml. with water.

10 ml. of this stock solution and 37.5 ml. perchloric acid (70 per cent.) are diluted to 500 ml. with ethanol.

(b) *Procedure*

In an Evelyn colorimeter tube are mixed 5 ml. ester solution and 1.5 ml. hydroxylamine solution. The mixture is allowed to stand at 25 °C. for 30

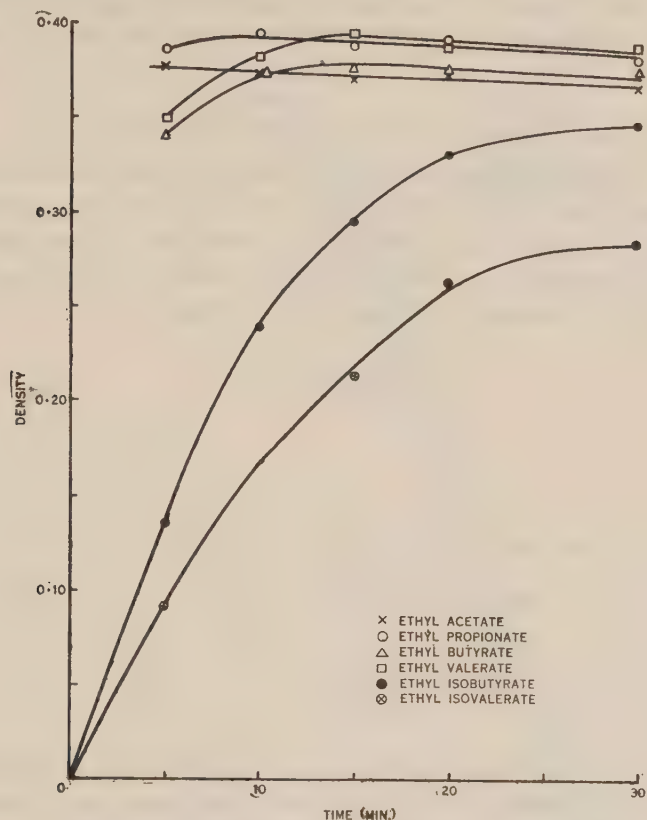


Fig. 2.—Formation of hydroxamic acids from different esters in ether at 25 °C. (Concentration of esters in ether  $8 \times 10^{-4}$  M.)

minutes and 12.5 ml. ferric perchlorate reagent solution are added. The solution is shaken and after 10 minutes compared in the Evelyn colorimeter with a similarly prepared blank, using the 520 filter.

#### IV. RATE OF FORMATION OF HYDROXAMIC ACIDS

The rates of reaction of a number of esters in ether, under the conditions of the Recommended Procedure, are shown in Figure 2. While the reactions of the straight carbon chain acids proceed at a fairly similar rate and are complete in 15 minutes, ethyl isobutyrate and ethyl isovalerate react much more slowly. The final colour is also less intense with the branched chain esters since there is time for considerable hydrolysis of the ester to take place before conversion to

hydroxamic acid is complete. The reactions of the ethyl esters of all the unbranched fatty acids containing 6-10 carbon atoms proceed at rates similar to the lower unbranched acids and are complete in 15 minutes. One series of three esters of the same acid was studied and ethyl, propyl, and amyl butyrates were found to react at the same rate.

Since some variability in the results was observed, when the reaction time was greater than 30 minutes and since approximately maximum formation of hydroxamic acid occurred in 30 minutes, this time was adopted as standard reaction time.

#### V. COMPARISON WITH THE COLOUR FROM PURE HYDROXAMIC ACIDS

The colours produced from ethyl acetate and ethyl propionate were compared with those produced from equivalent concentrations of pure acethydroxamic and propionhydroxamic acids, under the conditions of the Recommended Procedure. The results, which indicated approximately 85 per cent. conversion of esters to their hydroxamic acids, are given in Table 1.

TABLE 1

	Acetate	Propionate
Colour density from ethyl ester ( $8 \times 10^{-4}$ M) .. ..	0.365	0.380
Colour density from corresponding hydroxamic acid ( $8 \times 10^{-4}$ M) .. .. .	0.424	0.449
Degree of conversion of ester to hydroxamic acid ..	86%	84%

Acethydroxamic acid, m.p. 86-87 °C., was prepared from acetic anhydride(8) and from ethyl acetate(9, 10). Propionhydroxamic acid, m.p. 92-94 °C., was made from propionic anhydride(8). Acethydroxamic acid and propionhydroxamic acid were found to be quite stable under the conditions required for their formation from esters.

#### VI. ABSORPTION SPECTRA OF SOLUTIONS OF FERRIC-HYDROXAMATE COMPLEXES

The absorption spectra of solutions of ferric-hydroxamate complexes derived from pure hydroxamic acids and from esters (all prepared under the conditions of the Recommended Procedure) were obtained in cooperation with the Division of Physics, C.S.I.R.O., using a Hardy (General Electric) recording spectrophotometer. The absorption was expressed as density ( $2 - \log$  per cent. transmission) and was corrected by subtracting the density of the blank (Fig. 3). All the substances tested showed a maximum absorption of 520 mμ. Subsequent measurements with a Beckman quartz spectrophotometer were carried out on

the ethyl esters of the normal carboxylic acids containing 6–10 carbon atoms which also gave maxima at 520 m $\mu$ . The Evelyn filter 520 is obviously very suitable for the determination.

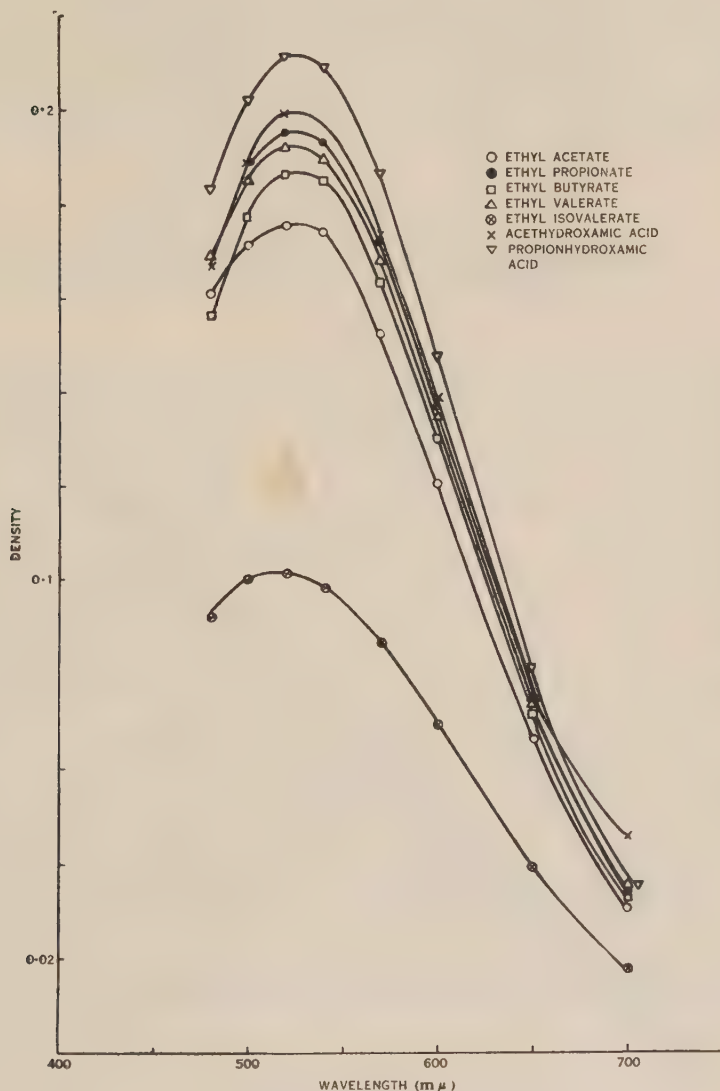


Fig. 3.—Spectral curves of ferric hydroxamate complexes from hydroxamic acids and esters. (Concentration of esters in ether  $8 \times 10^{-4}$  M.)

#### VII. RELATION BETWEEN COLOUR AND CONCENTRATION OF ESTER

Up to a  $1.6 \times 10^{-3}$  M concentration of ester in ether, the colour density is directly proportional to the concentration in accordance with Beer's law.



Calibration curves for some esters are shown in Figure 4. The ester concentration per unit density for various esters is given below :

Ethyl acetate	..	..	..	$2.14 \times 10^{-3}$ M
Ethyl propionate	..	..	..	$1.99 \times 10^{-3}$ M
Ethyl butyrate	..	..	..	$2.12 \times 10^{-3}$ M
Ethyl valerate	..	..	..	$2.07 \times 10^{-3}$ M
Ethyl hexanoate	..	..	..	$2.29 \times 10^{-3}$ M
Ethyl heptanoate	..	..	..	$2.08 \times 10^{-3}$ M
Ethyl octanoate	..	..	..	$2.04 \times 10^{-3}$ M
Ethyl nonanoate	..	..	..	$2.04 \times 10^{-3}$ M
Ethyl decanoate	..	..	..	$2.04 \times 10^{-3}$ M
Ethyl isobutyrate	..	..	..	$2.42 \times 10^{-3}$ M
Ethyl isovalerate	..	..	..	$2.91 \times 10^{-3}$ M

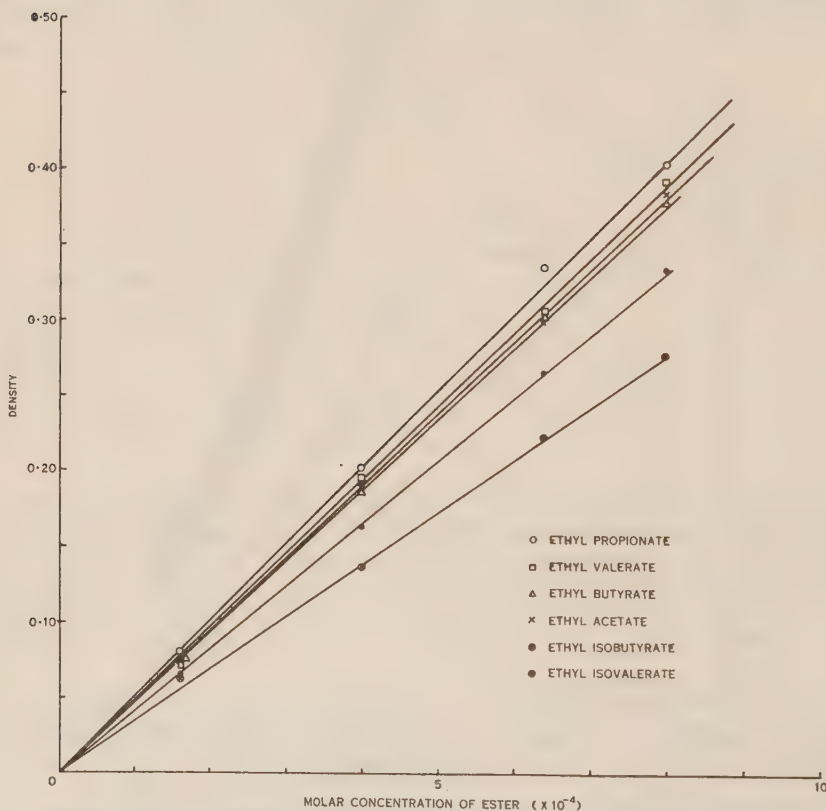


Fig. 4.—Relation between colour density and concentration of ester.

To calibrate a mixture of esters, one ester must be selected as arbitrary standard unless the mixture is simple and the relative amounts of each ester are known.

## VIII. COMPARISON WITH HILL'S METHOD

The Recommended Procedure was compared with the procedure of Hill(3, 4), which was not designed for volatile esters, in determinations of ethyl acetate, propionate, butyrate, and valerate.

The intensity of colour was very variable and weakest for ethyl acetate, which was, presumably, partly evaporated during the determination. With similar concentration of ethyl valerate, the iron hydroxamate formed in the method of Hill was about 73 per cent. of that with the method now adopted and the comparison was progressively worse with decreasing length of carbon chain and increasing volatility.

## IX. ACKNOWLEDGMENTS

The author wishes to thank Dr. F. E. Huelin for his suggestions and helpful criticism, and Miss Heather Smith for technical assistance.

## X. REFERENCES

- (1) ASSOCIATION OF OFFICIAL AGRICULTURAL CHEMISTS.—“Official and Tentative Methods of Analysis.” 6th Ed., pp. 194, 205. (Washington, 1945.)
- (2) FEIGL, F.—“Laboratory Manual of Spot Tests.” pp. 186-8. (Academic Press Inc.: New York, 1943.)
- (3) HILL, U. T.—Colorimetric determination of fatty acids and esters. *Ind. Eng. Chem. (Anal. Ed.)* **18**: 317-9 (1946).
- (4) HILL, U. T.—Colorimetric determination of fatty acids and esters. *Ind. Eng. Chem. (Anal. Ed.)* **19**: 932-3 (1947).
- (5) KEENAN, A. G.—Colorimetric determination of ethyl acetate. *Canad. Chem. Process Ind.* **29**: 857-8 (1945).
- (6) BAUER, F. C., and HIRSCH, E. F.—A new method for the colorimetric determination of the total esterified fatty acids in human sera. *Arch. Biochem.* **20**: 242-50 (1949).
- (7) REIMERS, F.—Investigations on the autoxidation of diethyl ether. Pts. 1-4. *Quart. J. Pharm.* **18**: 350-9 (1945); **19**: 27-38 (1946); **19**: 172-87 (1946); **19**: 473-9 (1946).
- (8) MIOLATI, A.—Ueber eine neue Bildungsweise der Hydroxamsäuren der Fettreihe. *Ber. dtsh. chem. Ges.* **25**: 699-701 (1892).
- (9) JONES, L. W., and SCOTT, A. W.—New hydroxamic acids derived from cyclopropane carboxylic acid, isobutyric acid and dibenzyl-acetic acid. A comparative study of the Beckmann rearrangement of their derivatives. *J. Amer. Chem. Soc.* **44**: 407-23 (1922).
- (10) INOUE, Y., and YUKAWA, H.—Separation and identification of aliphatic acids. Pt. 1. Hydroxamic acids derived from saturated aliphatic acids. *J. Agric. Chem. Soc. Japan* **16**: 504 (1940).

# STUDIES OF THE WALLACH TRANSFORMATION

## I. THE PRODUCTS OF THE REACTION

By P. H. GORE\* and G. K. HUGHES\*

[Manuscript received August 29, 1949]

### *Summary*

The products formed in the Wallach Transformation of various azoxybenzenes have been systematically examined. Many abnormal results claimed have been reinvestigated.

## I. INTRODUCTION

The conversion of azoxybenzenes by sulphuric acid to derivatives of 4-hydroxyazobenzene, known as the Wallach Transformation,<sup>†</sup> is usually accompanied by the formation of other products, namely, 2-hydroxyazobenzenes, azobenzenes, amorphous polymeric powders, sulphonic acids, and amines. Owing to the many conflicting statements in the literature concerning the products formed, their yields, and suggested mechanisms of formation, it was decided to study the transformation in detail. A considerable amount of previous work had to be repeated before a theoretical investigation of the problem could be attempted. Hence it was decided to confine Part I to the study of the types of product formed. Discussion of the mechanisms will be given in future papers. Each of the above products will be treated under a separate heading.

### *(a) 4-Hydroxyazobenzenes*

Wallach and Belli(1) (cf. Wallach and Kiepenheuer 2) first reported the isomerization of azoxybenzene to 4-hydroxyazobenzene, but Lachman(3), the first to study it in any detail, found that the yield of this product varied within wide limits (up to 70%), depending on many factors, principally concentration of the sulphuric acid solution (a maximum was found with monohydrate), ratio of the reactants, temperature, and time of reaction. Lukashevish and Kurdyumova(4) found a maximum of 61 per cent. with monohydrate at room temperature for 8½ days.

Experiments I–XVI recorded in this paper show a range of yield from 4 to 65 per cent. With 98 per cent. sulphuric acid at 20 °C., for instance, the maximum yield (28%) is obtained after 2 hours, when the rate of formation of 4-hydroxyazobenzene (represented by the ascending curve in Fig. 1) is just balanced by its rate of destruction (by sulphonation and polymerization; cf.

\* School of Chemistry, University of Sydney.

† Hereafter abbreviated to W.T.

Table 5; Expt. XXV; also Lachman 3). Its gradual removal from the reaction mixture is shown by the shape of the curve, which resembles a decaying exponential.

Whilst the ratio of reactants seems to play only a minor part, the effect of increased temperature is a higher yield of hydroxyazobenzene. Above 100 °C. the reaction tends to proceed with violence, with the yield of polymer and amine increasing rapidly (cf. Lachman 3).

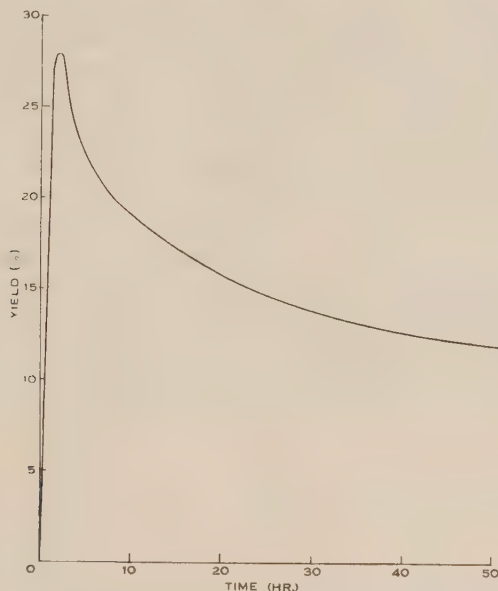


Fig. 1.—Yield of hydroxyazobenzene/time,  
98%  $\text{H}_2\text{SO}_4$ , 25 °C.

Figure 2 represents the influence of the concentration of the sulphuric acid on the yield of hydroxyazobenzene. It is seen that a maximum occurs with an acid corresponding to  $\text{H}_2\text{SO}_4 \cdot 1.1\text{H}_2\text{O}$  (83% w/w); with more dilute acid the reaction proceeds very slowly since solution of the azoxybenzene is only gradual. (With 75% sulphuric acid ( $\text{H}_2\text{SO}_4 \cdot 1.84\text{H}_2\text{O}$ ) the rearrangement gives no azobenzene or polymer and 88 per cent. recovered azoxybenzene.) The net yield of hydroxyazobenzene (in the quoted instance 84%) is given by the dotted line of Figure 2.

The rearrangements of many other azoxy compounds have given the expected product, but some very surprising and anomalous results have been claimed.\* It has been found impracticable to include all the examples, and only those which were regarded as extraordinary are discussed.†

\* An example is Klinger and Pitschke's(5) claim of having detected some formic acid from a W.T. with 3,3'-dinitroazoxybenzene, although formic acid was then known to dehydrate under the reaction conditions(6). The explanation lies in the fact that their method of preparation gives an azoxy compound heavily contaminated with sodium formate.

† It is hoped, at the conclusion of these studies, to write a review on this reaction.



According to Klinger and Pitschke(5) treatment of 2,2'-dimethylazoxybenzene with concentrated sulphuric acid at 100–120 °C. yielded 2,2'-dimethylazobenzene and amorphous material, but no hydroxyazotoluene. This observation caused some confusion and delayed systematic investigation considerably (cf. Knipscheer 7; Parsons and Bailar 8). However, by using monohydrate (85 % w/w) and conducting the reaction at room temperature, a 40 per cent. yield of 2,2'-dimethyl-4-hydroxyazobenzene is obtained.

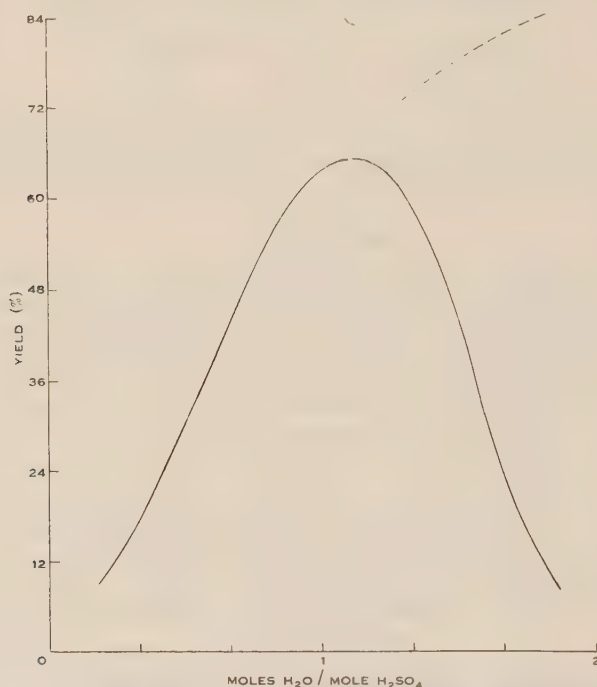


Fig. 2.—Yield of hydroxyazobenzene/hydration of acid, 90 °C.,  $\frac{1}{2}$  hr.

Klinger and Pitschke(5) also treated 3,3'-dinitroazoxybenzene with concentrated sulphuric acid at 140°C. and claimed a nearly quantitative conversion to 4-hydroxy-3,3'-dinitroazobenzene. Using the conditions given above, a maximum yield of 56 per cent. is obtained, accompanied by the usual by-products. No explanation is offered for this discrepancy.

Rotarski(9) treated 4,4'-dimethoxyazoxybenzene with sulphuric acid but isolated only resinous material; however, besides a high yield of polymer, about 0.6 per cent. of 4-hydroxy-4'-methoxyazobenzene is formed. It should be noted that under similar conditions 4,4'-dimethoxyazobenzene gives no phenolic compound (Expt. XXVI).

#### (b) 2-Hydroxyazobenzenes

Bamberger(10) first observed the formation of 2-hydroxyazobenzene from azoxybenzene, quoting a yield of 0.6 per cent. Knipscheer(7) stated that its

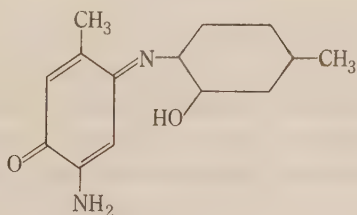
yield increased markedly with rise in temperature, but gave no supporting experimental evidence. In some cases hydroxyazo compounds have been reported in the rearrangement of 4,4'-disubstituted azoxybenzenes(11-14), and these were undoubtedly derivatives of 2-hydroxyazobenzene, though in no case was a structure proposed.

In an obscure footnote Bamberger(15) mentioned that on treatment of 4,4'-dimethylazoxybenzene, with sulphuric acid at 20-25°C., a product, m.p. 204°C., was isolated which was not steam volatile. Parsons and Bailar(8), apparently unaware of this previous work, reported the isolation of an alkali soluble (*A*) and an alkali insoluble (*B*) fraction from this reaction, both of which analysed for  $C_{14}H_{14}ON_2$ . They synthesized 2-ethoxy-4,4'-dimethylazobenzene (m.p. 76 °C.) which depressed the melting point of the "ethyl ether" of *A* (m.p. 50 °C.), and suggested the structure of *A* to be 3-hydroxy-4,4'-dimethylazobenzene. Since *B* also gave a benzoate its structure was claimed to be 4-methyl-4'-hydroxymethylazobenzene.

The striking similarity in properties of *A* and *B*, apart from the solubility in alkali, seemed more than coincidental and no depression of a mixed melting point shows their identity. It appears that the sodium salt of the substance is only slightly soluble especially in caustic soda solutions of more than 1 per cent. It must be 2-hydroxy-4,4'-dimethylazobenzene, which is represented by



as on reduction it gives toluidine and an aminophenol, which oxidizes to 2-methyl-5-amino-*p*-benzoquinone 1-(2'-hydroxy-4'-methyl anil)(16, 17).



If on ethylation of the hydroxyazotoluene a large proportion of the *N*-ethyl derivative (m.p. 50 °C.) and only a small amount of the *o*-ethyl isomer(m.p. 76 °C.) were formed, then the abovementioned melting point depression would be explained. A "tolyl" azo-*m*-cresol, synthesized by McPherson and Boord(18), was most probably 2-hydroxy-2',4-dimethylazobenzene, and not 2-hydroxy-4,4'-dimethylazobenzene as suggested by Parsons and Bailar(8).

Apparently the substitution of methyl groups in the benzene rings of 2-hydroxyazobenzene lowers the solubility of its sodium salt in alkali (cf. the above with 2-hydroxyazobenzene(10) and 2-hydroxy-3,5,6,2',4',5'-hexamethylazobenzene(19)).

This rearrangement of 4,4'-dimethylazoxybenzene constitutes the first case on record of the W.T. giving a 2-hydroxyazobenzene derivative in good yield, i.e. 47 per cent.

(c) *Azobenzenes*

No plausible explanation has been offered for the formation of azobenzene in the W.T. from the time of its isolation by Wallach and Belli(1), but it was recognized as a usual by-product by Lachman(3), who stated that under certain conditions the yield reached 60 per cent.

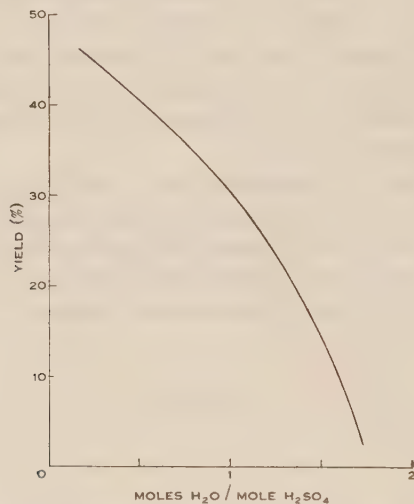


Fig. 3.—Yield of azobenzene/hydration of acid,  
90 °C.,  $\frac{1}{2}$  hr.

From a large number of experiments it is seen that the formation of azobenzene is favoured by a high concentration of sulphuric acid, which is favourable also to the formation of the polymer (see below).

Figure 3 shows the effect of concentration of acid on the yield of azobenzene. It is seen that the yield falls off rapidly with increased dilution of the sulphuric acid. No azobenzene is isolated from a reaction with 75 per cent. sulphuric acid.

Lachman(3) isolated a 15 per cent. yield of azobenzene from a heated mixture of 4-hydroxyazobenzene and monohydrate and concluded that the azobenzene produced during the W.T. occurs as the result of a secondary action of sulphuric acid upon the rearrangement product 4-hydroxyazobenzene. His words are: "It seems clear that the formation of azobenzene is the result of auto-reduction of hydroxyazobenzene." This view is untenable for at least two reasons: (i) with a more concentrated acid, a considerable increase in the proportion of azobenzene is to be expected, whereas analysis of a reaction mixture of 4-hydroxyazobenzene and 98 per cent. sulphuric acid gives only about 3 per cent. azobenzene (Expt. XXV); (ii) on prolonged action of sulphuric

acid it would be expected that the proportion of azobenzene, which sulphonates at a very slow rate (Expt. XXVII), would increase steadily at the expense of 4-hydroxyazobenzene, but this is not so, as examination of Figure 4 indicates.

It is difficult to see any correlation between the percentage yield of the derived azo compounds and the structure of the azoxy compound. Yields vary from traces, as for 3,3'-dichloroazoxybenzene(13) or 2,2'-dimethoxyazoxybenzene (Expts. XX-XXI) to nearly quantitative conversion as claimed for 4,4'-dichloroazoxybenzene(13, 20).\*

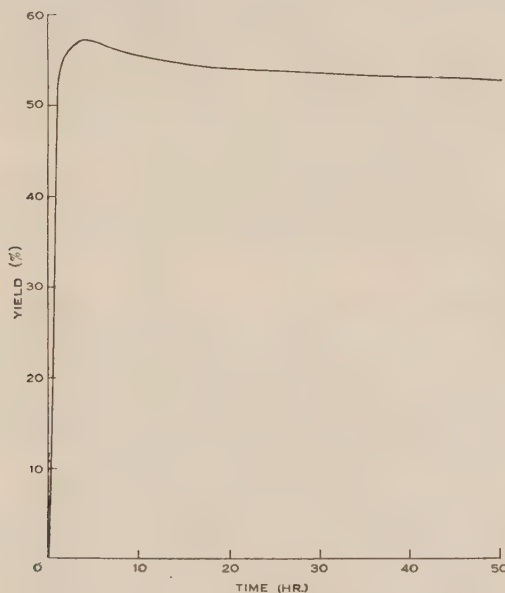


Fig. 4.—Yield of azobenzene/time,  
98%  $\text{H}_2\text{SO}_4$ , 25 °C.

Bamberger(21) and Angeli(22) believed that the azo compounds were formed by direct reduction.

In the cases investigated the yield of azo compound varies between unweighable traces and 55 per cent.

Previous workers(8, 15) failed to isolate 4,4'-dimethylazobenzene from the reaction with 4,4'-dimethylazoxybenzene, but a 22 per cent. yield of this azo compound is obtained by the use of a systematic method of separation (Expt. XXIV).

#### (d) Polymer

The polymer formed in the reaction with azoxybenzene was fully described by Lachman(3) who regarded it as the oxidation product of hydroxyazobenzene, ( $x \text{ C}_6\text{H}_5\text{.N:N.C}_6\text{H}_4\text{OH.(n-x)O}$ ).

\* This will be discussed in a later paper.



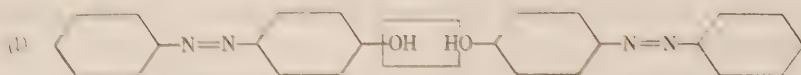
The colour of the polymer (yields obtained, 1.6–21%) ranges from medium brown (reaction time, 1 hour) to jet black (reaction time, 3 months). The polymer is slightly soluble in alcohol (more so the "young" polymer), completely soluble in pyridine and 10 per cent. sodium hydroxide solution, appreciably soluble in sodium carbonate solution (5N), aqueous ammonia (sp. gr. 0.880), and aniline and thus is acidic. The colour of the above solutions is an intense red-brown, suggesting the presence of azo groups.

A typical analysis is C, 61.6; H, 4.1; N, 10.6; S, 3.3% (O, 20.4%), which approximates to an (polymeric) azo compound and is quite far from "free carbon". Since the polymer is acidic, and contains sulphur, it can reasonably be assumed that it contains sulphonic acid groups. It would be expected that a polymer would increase in sulphonic acid content (and thus in % S) with time; this is borne out by analytical results (polymer, reaction time 1 week: S, 3.6%; reaction time 3 months: S, 4.2%).

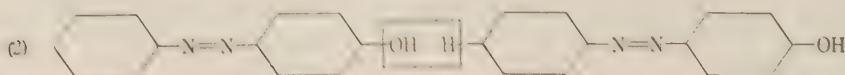
Lachman's(3) observation, that "aniline" (i.e. aniline, see below) is split off from the polymer by treatment with various alkaline or acidic agents, is probably best explained by considering some aniline to be adsorbed on the polymeric molecules.

In order to elucidate the structure of the W.T. polymer, it is necessary to consider the linkages present. Of the oxygen content of the polymer 4.9 per cent. is accounted for as sulphonic acid groups, another 3.7 per cent. must occur as hydroxyl groups (from the methoxy content of the polymer after methylation: 7.1%), the remainder (11.8%) would occur as ether linkages, together with some methylation resistant hydroxyl groups.

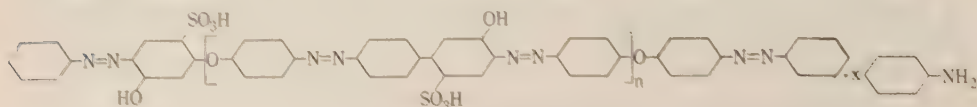
By treatment of the normal rearrangement product with monohydrate Lachman(3) obtained a 1 per cent. yield of polymer. Using 98 per cent. sulphuric acid only a small trace of polymer is isolated. Apparently, therefore, the main sources of the polymer is not 4-hydroxyazobenzene (cf. comments on the formation of azobenzene), however, from which polymerization can take place by dehydration of 2 types:



giving an "ether" linkage, or



giving a C-C linkage, and it is probable that linkages of both types occur in the W.T. polymer, which is best represented by



with a variable  $n$ , and a varying number of hydroxyl and sulphonic acid groups, and possibly some *ortho* "ether" and C-C linkages.

The yield of polymer is largely dependent on the strength of the sulphuric acid; Figure 5 suggests a dehydration type of mechanism. It is seen that whenever conditions favour dehydration, such as high concentration of acid or temperatures above 100 °C. (cf. Lachman 3), the yield of polymer is correspondingly increased. In the case of the yield of polymer increasing with time of the reaction, i.e. after complete removal of azoxybenzene, the source clearly is hydroxyazobenzene (cf. Fig. 1) (also possibly sulphonic acids—see below), the yield of which decreases with time.

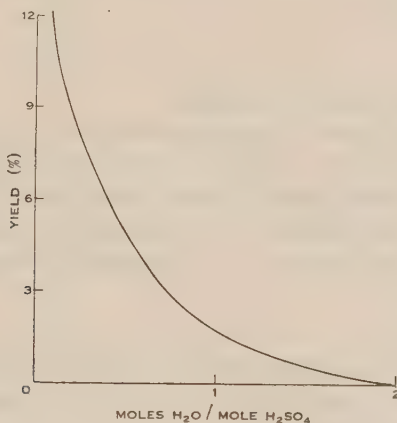


Fig. 5.—Yield of polymer/hydration of acid, 90 °C.,  $\frac{1}{2}$  hr.

Before discussing the factors controlling polymer formation from substituted azoxybenzenes it is necessary to mention a number of cases in the literature in which either the polymeric nature of the product was not recognized or the constitution of the material isolated was misunderstood.

#### 1,4-Bisbenzeneazoxybenzene(22),



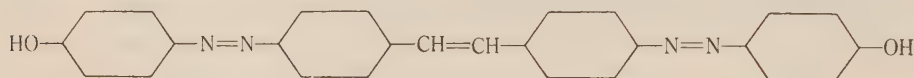
on treatment with concentrated sulphuric acid was easily transformed into two alkali soluble products, the one soluble in benzene being identified as 4'-hydroxy-1,4-bisbenzeneazobenzene,



the other being so insoluble in all solvents as to make it impossible to characterize it, but its structure was given as



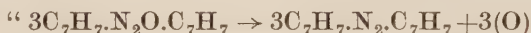
Later Bigiavi and Sabatelli(23), on treatment of  $\alpha$ -4-methylazoxybenzene with sulphuric acid, found besides the normal rearrangement product a benzene insoluble black powder, m.p.  $>250^\circ\text{C}$ ., N, 11.7%, easily soluble in sodium hydroxide solution giving a yellow-brown solution and suggested its structure as



and quoted the previous case in support.

It seems certain that each of these products was a typical W.T. polymer.

Klinger and Pitschke(5) obtained from 2,2'-dimethylazoxybenzene a mixture of amorphous dark coloured acids, part only being soluble in alcohol. The structures of the two fractions were suggested to be 2-methylazobenzene-2'-carboxylic acid and azobenzene 2,2'-dicarboxylic acid, and as mechanism was suggested,



The findings about the amorphous substances obtained in this reaction on re-examination are given below.

Further, in a number of cases dark amorphous materials were obtained but not recognized as polymers (cf.  $\alpha$ - and  $\beta$ -2-hydroxyazoxybenzene(21),  $\alpha$ -azoxynaphthalene(24), and 3,3'-dinitroazoxybenzene(5).

A summary is given in Table 1 of some of the yields and solubilities in

TABLE 1

Azoxybenzene			Yield of Polymer (%)	Solubility in 10% Aqueous NaOH
2,2'-Dimethoxy	..	..	38	Very slightly
4,4'-Dimethoxy	..	..	34	Easily
4,4'-Dimethyl	..	..	23	Slightly
3,3'-Dinitro	..	..	8.4	Easily
2,2'-Dimethyl	..	..	5.1	Easily

10 per cent. sodium hydroxide solution of polymers derived from the azoxy compounds examined. Although these polymers are not obtained under standard conditions (yields therefore are not strictly comparable) it can be

seen that a high yield of polymer is obtained in those instances where the azoxy compound contains electron-donating substituent groups (e.g. methoxy groups), and a low yield when the azoxy compound contains electron-attracting substituent groups (e.g. nitro groups). As may be expected a high yield of polymer is generally accompanied by a high yield of sulphonic acids.

The polymer derived from 4,4'-dimethoxyazoxybenzene contains about 1-3 per cent. methoxy groups; assuming that there are only two methoxy groups per polymer molecule, one at each end, the mean molecular weight should be 2000-7000. Using a similar method of reasoning to that used with the polymer derived from azoxybenzene, a specimen polymer from

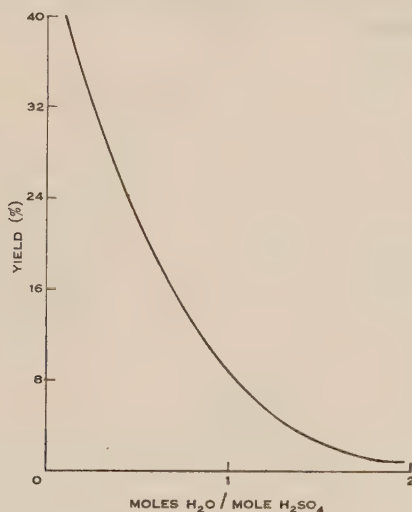


Fig. 6.—Yield of sulphonic acids/hydration of acid,  
90 °C.,  $\frac{1}{2}$  hr.

4,4'-dimethoxyazoxybenzene (Analysis: C, 56.4; H, 3.9; N, 9.6; S, 3.4; -OMe, 3.1%; methylated polymer: -OMe, 12.9%) has 5.1 per cent. oxygen as sulphonic acid groups, 5.0 per cent. oxygen as hydroxyl groups, 1.6 per cent. oxygen as methoxy groups and 15.0 per cent. oxygen in ether linkages and/or methylation resistant hydroxyl groups.

This polymer separated into alcohol soluble (18%) and alcohol insoluble (82%) fractions; by comparison of their sulphur (3.5%, 2.5% resp.) and methoxy (2.5%, 0.9% resp.) analysis figures and mol. wt. estimations (2500, 6900 resp.) (based on the methoxy content) it is seen that a high proportion of the sulphonic acid residues aggregate at the ends of the polymeric molecule.

The polymer derived from 2,2'-dimethylazoxybenzene separated into approximately equal weights of alcohol soluble and insoluble fractions.

It is observed that a methylated polymer, i.e. one in which the hydroxyl groups have been replaced by methoxy groups, is only sparingly soluble in 10 per cent. sodium hydroxide solution. This suggests that the solubility in alkali of a polymer depends not so much on its sulphonic acid content as on its content of hydroxyl groups.



One observation which may have a bearing on the formation of polymer is its slow precipitation from the clear mother liquor from the W.T. of 3,3'-dinitro-azoxybenzene. This seems to suggest that polymerization may take place via the sulphonic acid.

(e) *Sulphonic Acids*

From the mother liquor of the reaction between azoxybenzene and sulphuric acid, Lachman(3) isolated 4-hydroxyazobenzene-4'-sulphonic acid. Another sulphonic acid probably present would be 2-hydroxyazobenzene-4'-sulphonic acid. Whilst Lachman actually isolated the sulphonic acid as its sodium salt, in the present work this method was found to be too laborious, and yields quoted were obtained by difference.

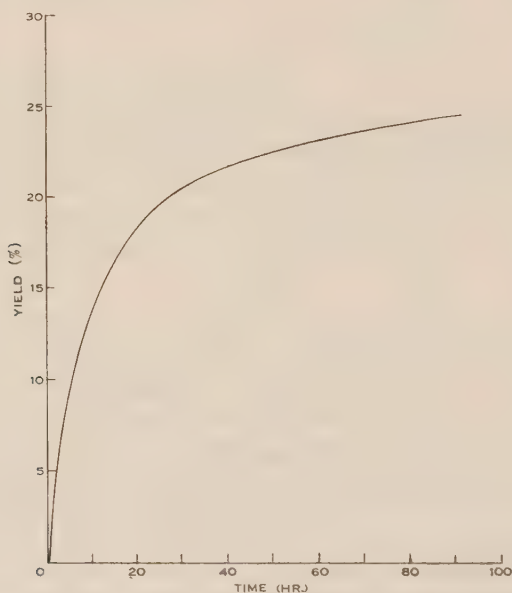
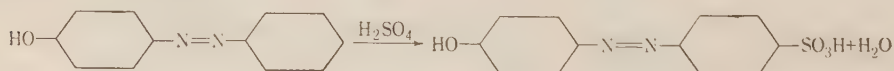


Fig. 7.—Yield of sulphonic acids/time,  
98%  $\text{H}_2\text{SO}_4$ , 25 °C.

The yield of sulphonic acids depends on the same factors that govern the formation of polymer (cf. Figs. 5 and 6). This is in accordance with the possible formation of polymer via the sulphonic acid.

Lachman(3) suggested that 4-hydroxyazobenzene-4'-sulphonic acid is formed by direct sulphonation of 4-hydroxyazobenzene,



From the shape of the curve in Figure 7 it can be seen that the yield of sulphonic acids depends at first on the amount of azoxybenzene present in the solution; the rate of its formation gradually falls off, though the quantity of 4-hydroxyazobenzene has reached a maximum. The yield of sulphonic acids still increases,

but only at a slow rate, when no azoxybenzene is left in the solution, and its mode of formation is then best given by Lachman's(3) idea. It appears then that the formation of 4-hydroxyazobenzene-4'-sulphonic acid directly from azoxybenzene is also a possibility.

A comparison of the yields of soluble sulphonic acids from the various substituted azoxybenzenes examined (Table 2) shows that a high yield is obtained with an azoxy compound which would be expected to be readily sulphonated. The anomaly of 3,3'-dinitroazoxybenzene is only an apparent one, since its rearrangement was effected under much more vigorous conditions than in the other cases.

TABLE 2

Azoxybenzene	Yield (%)
2,2'-Dimethoxy .. ..	56
4,4'-Dimethoxy .. ..	30-37
3,3'-Dinitro .. ..	23
4,4'-Dimethyl .. ..	7-8
2,2'-Dimethyl .. ..	1-2

#### (f) Amines

Lachman(3) isolated a primary amine from the W.T. by steam distillation of the basified mother liquor, which he described as *n*-amylamine. As might be expected, the base is, in fact, aniline. It was definitely identified by melting point comparisons of the hydrochloride and 5-bromosalicylidene derivative(25) with authentic specimens.

The general conditions favouring formation of this amine were stated by Lachman(3); e.g. by conducting a reaction between azoxybenzene and twice its weight of concentrated sulphuric acid at a high temperature, a violent reaction took place and a 14 per cent. yield (w/w) of amine was obtained. At normal temperatures 0.3-0.7 per cent. and at 100 °C. about 1.6 per cent. of aniline are usually obtained. This increase in yield with rise in temperature is probably related to the formation of aniline from azoxybenzene by the action of heat(26, 27).

Of the other compounds investigated, only 2,2'-dimethoxy-, 4,4'-dimethoxy-, and 2,2'-dimethylazoxybenzenes were examined for amines; these substances were found to give *o*-anisidine (5.8%), *p*-anisidine (2.3%), and *o*-toluidine (0.8%) respectively.

## II. EXPERIMENTAL

### (a) General Method of Separation

The solution of azoxy compound in sulphuric acid was poured into ice water, the precipitated solid collected by suction on a sintered glass funnel, washed twice with water, and drained. The residue was exhaustively extracted with ether, which dissolved out the hydroxyazo, azo, and

any unchanged azoxy compounds. This ether solution (together with the ether extract of the mother liquor), after extraction with 10% sodium hydroxide solution and removal of the solvent gave the azo and any azoxy compound. The proportion of the latter in the mixture was determined by m.p.

The alkaline extract on acidification with dilute acetic (or hydrochloric) acid precipitated the hydroxyazo compound(s).

The amorphous residue was washed with water, then dried at 110 °C. On steam distillation of the aqueous mother liquor made alkaline with sodium hydroxide, the amine was collected in hydrochloric acid, followed by evaporation to dryness.

The steam distillation residue still contained sulphonic acids, the yield of which was determined by difference.

(b) *Azoxybenzene*

Yields obtained are set out in Tables 3-5.

TABLE 3  
USING 98% SULPHURIC ACID

Expt.	Conditions	Hydroxyazo- benzene (%)	Azo- benzene (%)	Polymer (%)	Aniline (%)	Sulphonic Acids (%)
I	6.5 g., 41 ml.; 25 °C., 187 hr.	9.5	58	14	0.5	18-19
II	2 g., 2 ml.; 95 °C., 5.6 hr.	46	35	13	1.6	4-5

TABLE 4  
USING AZOXYBENZENE (5 G.): SULPHURIC ACID (50 ML.), 90 °C., 30 MIN.

Expt.	Concen- tration (%)	Hydration	Hydroxy- azobenzene (%)	Azo- benzene (%)	Azoxy- benzene (%)	Polymer (%)	Aniline and Sul- phonic Acids (%)
III	67.5	2.63			100		
IV	75	1.84	9.8	0	88	Trace	2
V	83	1.11	65	29	Trace	1.9	4
VI	85	0.96	62	29	0	1.6	8
VII	90	0.64	36	39		3.7	21
VIII	98	0.11	7.8	46		11	36

The accuracy of these determinations was  $\pm 3\%$ .

Whilst the hydroxyazobenzene formed in Expts. I-VIII, and XII-XVI was practically pure (m.p. c. 150 °C.), the product in Expts. IX-XI (m.p. c. 128-134 °C.) was apparently mixed with some other phenolic material.

TABLE 5  
USING AZOXYBENZENE (5 G.): 98% SULPHURIC ACID (50 ML.), 25 °C.

Expt.	Time	Hydroxy-azobenzene (%)	Azobenzene (%)	Azoxybenzene (%)	Polymer (%)	Aniline and Sulphonic Acids (%)
IX	1 hr.	21	51	12	14	2
X	2 hr.	28	55	Trace	11	6
XI	4 hr.	23	57	0	13	7
XII	22 hr.	16	54		11	18
XIII	49 hr.	15	53		14	18
XIV	73 hr.	12	52		14	23
XV	97 hr.	8.4	53		14	24
XVI	3 months	4.4	53		21	22

*Analyses of Polymers.*—I (Found: C, 62.0; H, 3.8; N, 10.0; S, 3.6%). IX (Found: C, 65.0; H, 3.9; N, 10.4%). XII (Found: N, 10.9%). XIII (Found: N, 10.2%). XIV (Found: N, 10.4%). XV (Found: N, 10.9%). XVI (Found: S, 4.2%). Mixed sample M (Found: C, 61.6; H, 4.1; N, 10.6; S, 3.3%). Methylated M (Found: -OMe, 7.1%). The yields of polymer were based on the mean nitrogen content.

The aniline formed in this way gave a hydrochloride, m.p. 186–188 °C. (crude), 193 °C. (recrystallized), and a 5-bromosalicylidene derivative (25), m.p. 121 °C.; both derivatives were undepressed in m.p. by admixture with authentic specimens.

The colour of the steam distillation residue was red.

*Methylation of Polymers.*—The polymer was dissolved in an excess warm 10% sodium hydroxide solution, and dimethyl sulphate added in small portions, until no further precipitate was formed. This was followed by two further additions of sodium hydroxide solution and dimethyl sulphate. The reaction was completed by heating on the water-bath for  $\frac{1}{2}$  hour. The methylated polymer is collected on a sintered glass funnel, washed, and dried at 110 °C.

(c) 2,2'-Dimethylazobenzene

The azoxy compound (10 g.): 85% sulphuric acid (100 g.) was kept at 25 °C. for 120 hours (the solution warmed to 35 °C. during the first hour).

TABLE 6

Expt.	Hydroxyazo Compound	Azo Compound	Polymer	<i>o</i> -Toluidine	Sulphonic Acids
XVII	40%	53%	5.2%	0.8%	1–2%

The phenol had m.p. 108 °C., or 111 °C. after recrystallization from benzene or dilute acetic acid, and was thus 4-hydroxy-2,2'-dimethylazobenzene(28).



*4-Benzoyloxy-2,2'-Dimethylazobenzene*.—This was prepared by a Schotten-Baumann reaction. Orange needles from dilute acetic acid, m.p. 98 °C. (Found: C, 75.9; H, 5.5; N, 8.4%. Calculated for  $C_{21}H_{18}O_2N_2$ : C, 76.3; H, 5.5; N, 8.5%.)

Polymer XVII (Found: N, 7.1%) separated into alcohol soluble (Found: C, 79.5; H, 6.8; N, 11.6%) and insoluble (Found: C, 64.7; H, 4.5; N, 5.8; S, 1.3%) fractions (approx. equal weights).

The amine gave a 5-bromosalicylidene derivative, m.p. 85 °C., undepressed in m.p. by admixture with authentic 5-bromosalicylidene-*o*-toluidine(29).

The colour of the steam distillation residue was brown.

(d) *4,4'-Dimethoxyazoxybenzene*

TABLE 7  
USING 85% SULPHURIC ACID

Expt.	Conditions	Hydroxy- azo Compound (%)	Azo Compound (%)	Polymer (%)	Amine (%)	Sulphonic Acids (%)
XVIII	5 g., 40 ml.; 25 °C., 40 hr.	0.6	26	35	2.3	36-37
XIX	1.1 g., 10 ml.; 90 °C., 18 min.	Trace	33	36		

The hydroxyazo compound formed pink needles, m.p. 134-136 °C. (crude) or 140 °C. (pure), and did not depress the m.p. of authentic 4-hydroxy-4'-methoxyazobenzene(30). The benzoyl derivative, m.p. 161 °C. (clearing point 178 °C.) did not depress the m.p. of authentic 4-benzoyloxy-4'-methoxyazobenzene (reported clearing point 173 °C.(31)).

Polymer XVIII (Found: C, 56.4; H, 3.9; N, 9.6; -OMe, 3.1; S, 3.4%) had mean mol. wt. = wt. of 200(-OMe)/%(-OMe)  $\div$  2000. Methylated Polymer XVIII (Found: N, 7.1; -OMe, 12.9%). Polymer XIX separated into an alcohol soluble fraction (18%) (Found: N, 7.1; -OMe, 2.5; S, 3.5%), mean mol. wt.  $\div$  2500, and alcohol insoluble fraction (82%) (Found: N, 5.5; -OMe, 0.9; S, 2.5%), mean mol. wt.  $\div$  6900.

The amine gave a 5-bromosalicylidene derivative, m.p. 155-156 °C., which did not depress the m.p. (156 °C.) of 5-bromosalicylidene-*p*-anisidine(29).

The colour of the steam distillation residue was a very dark burgundy red, indicating a large quantity of sulphonic acids. Removal of sulphonic acid groups was attempted in two ways: (i) refluxing the diluted mother liquor for 5-6 hours, and (ii) passing superheated steam into the mother liquor during 2 hours. Ether extraction resulted in pale extracts leaving only traces of 4,4'-dimethoxyazobenzene on evaporation.

(e) *2,2'-Dimethoxyazoxybenzene*

Polymer XXI (Found: N, 6.0; -OMe, 11.8; S, 3.8%) was only very slightly soluble in 10% sodium hydroxide solution.

The amine gave a 5-bromosalicylidene derivative, m.p. 109 °C., undepressed in m.p. by admixture with 5-bromosalicylidene *o*-anisidine (m.p. 110 °C.(25)).

The mother liquor was a dark purplish red.

TABLE 8  
 USING 85% SULPHURIC ACID

Expt.	Conditions	Hydroxy- azo Compound (%)	Azo Compound	Azoxy Compound (%)	Polymer (%)	Amine (%)	Sulphonic Acids (%)
XX	2.2 g., 30 ml. ; 20 °C., 30 min.	Trace	0	82	Trace	Trace	17-18
XXI	5 g., 50 ml. ; 90 °C., 10 min.	0.2	Trace	0	38	5.8	56

## (f) 3,3'-Dinitroazoxybenzene

TABLE 9

Expt.	Conditions	Hydroxyazo Compound (%)	Azo Compound (%)	Polymer (%)	Amine and Sulphonic Acids (%)
XXII	5 g., 50 ml. 85% ; heated from 20- 220 °C., over 2 hr.	Trace	Trace	14	86
XXIII	5 g., 70 ml. 98% ; 140 °C., 1.8 hr.	56	2.8	(1) 8.4 (2) 9.9	23

The hydroxyazo compound formed yellow crystals from benzene, orange from dilute acetic acid, m.p. 176-177 °C. (m.p. for 4-hydroxy-3,3'-dinitroazobenzene is 172-173 °C.(5), or 179 °C.(32)). The phenol was acetylated in pyridine, using acetic anhydride, at reflux temperature. The ester formed orange crystals from aqueous acetic acid, m.p. 134-136 °C. (m.p. 138 °C.(32)) (Found : N, 17.1%. Calculated for  $C_{14}H_{10}O_6N_4$  : N, 17.0%).

The "azo compound", soluble in ether and insoluble in sodium hydroxide solution, consisted of a brownish-red solid, m.p. 193-205 °C., and so was not identical with 3,3'-dinitroazobenzene (m.p. 153 °C.(33)). This product could not be purified further.

In Expt. XXII large scale destruction occurred, resulting in an almost black mother liquor ; Polymer XXII (Found : C, 50.5 ; H, 3.2 ; N, 5.4 ; S, 4.0%).

Polymer XXIII (1) (Found : C, 50.1 ; H, 2.7 ; N, 15.8%) was easily soluble in warm sodium hydroxide solution giving a red-brown solution. From the brown aqueous mother liquor a second polymer (2) separated over a period of 2 months, which was also easily soluble in alkali ; its nature was not otherwise investigated.

## (g) 4,4'-Dimethylazoxybenzene

TABLE 10

Expt.	Conditions	Hydroxyazo Compound	Azo Compound	Polymer	Amine and Sulphonic Acids
XXIV	10 g., 100 ml. 85% ; 90 °C., 30 min.	47%	22%	23%	7-8%

The ether extract (see Section II (a)) gave a mixture of 4,4'-dimethylazobenzene and 2-hydroxy-4,4'-dimethylazobenzene (see below) which were separated either by fractional crystallization from ligroin (the latter compound being obtained pure after 4 crystallizations), or better, by repeated extraction with 1% sodium hydroxide solution which finally gave pure 4,4'-dimethylazobenzene, m.p. 144-145 °C. (m.p. 144 °C.(34)), obtained from dilute acetic acid or aqueous ethanol as orange crystals.

*2-Hydroxy-4,4'-Dimethylazobenzene.*—This crystallized in orange-red needles, m.p. 210-212 °C. (Found: N, 12.2%. Calculated for  $C_{14}H_{14}ON_2$ : N, 12.4%.) On cooling a hot saturated (orange) solution of this substance in 10% sodium hydroxide, small orange crystals, m.p. 208-209 °C., of the sodium derivative settled out, which did not depress the m.p. of the phenol, but left an alkaline ash on ignition. It was ether insoluble and therefore did not contain any hydroxyazo compound.

*2-Benzoyloxy-4,4'-Dimethylazobenzene.*—This was prepared in pyridine solution using benzoyl chloride; crystallized from aqueous acetic acid, m.p. 122 °C. (Found: N, 8.5%. Calculated for  $C_{21}H_{18}O_2N_2$ : N, 8.5%.)

Stannous chloride-hydrochloric acid smoothly reduced the hydroxyazo compound to *p*-toluidine and an aminophenol, which was oxidized by aeration in aqueous solution to a brown product, m.p. 163-180 °C. (crude), 185 °C. (crystallized from benzene); the oxidation product of 3-hydroxy-4-aminotoluene, 2-methyl-5-amino-1,4-benzoquinone-1-(2'-hydroxy-4'-methyl anil) has m.p. 185-186 °C.(16, 17).

Polymer XXIV (Found: C, 64.4; H, 3.5; N, 8.9; S, 3.3%) was slightly soluble in dilute sodium hydroxide solution.

## (h) Action of Sulphuric Acid on

## (1) 4-Hydroxyazobenzene

Separation as for a W.T.

Expt.	Conditions	Recovered Hydroxy-azobenzene	Azobenzene	Polymer	Sulphonic Acids
XXV	2.00 g., 20 ml. 98% ; 25-30 °C., 23 hr.	48%	3%	1%	48%

## (2) 4,4'-Dimethoxyazobenzene

Expt.	Conditions	Unchanged Azo Compound
XXVI	0.95 g., 20 ml. 85% ; 20–25 °C., 18 hr.	99%

The diluted reaction mixture gave an ether extract, from which no phenol was isolated.

## (3) Azobenzene

Expt.	Conditions	Unchanged Azo Compound
XXVII	2.5 g., 50 ml. 98% ; 20–25 °C., 67 hr.	>99%

## III. ACKNOWLEDGMENTS

The authors' thanks are due to Miss J. Fildes and Mrs. E. Smith, microanalysts, University of Sydney, for the semi-microanalyses recorded in this paper, and to the Commonwealth Research Grant Committee, University of Sydney, for a scholarship awarded to one of them (P.H.G.), during the tenure of which this work was carried out.

## IV. REFERENCES

- (1) WALLACH, O., and BELLI, L.—*Ber. dtsh. chem. Ges.* **13** : 525 (1880).
- (2) WALLACH, O., and KIEPENHEUER, L.—*Ber. dtsh. chem. Ges.* **14** : 2617 (1881).
- (3) LACHMAN, A.—*J. Amer. Chem. Soc.* **24** : 1181 (1902).
- (4) LUKASHEVISH, V. O., and KURDYUMOVA, T. N.—*Chem. Abstr.* **43** : 3800 (1949).
- (5) KLINGER, H., and PITSCHE, R.—*Ber. dtsh. chem. Ges.* **18** : 2552 (1885).
- (6) DÖBEREINER, J. W.—*Liebigs Ann.* **17** : 73 (1836).
- (7) KNIPSCHER, H. M.—*Jaarb. Akad. Amst.* **5** : 51 (1902).
- (8) PARSONS, T., and BAILAR, J. C.—*J. Amer. Chem. Soc.* **58** : 268 (1936).
- (9) ROTARSKI, T.—*Ber. dtsh. chem. Ges.* **41** : 865 (1908).
- (10) BAMBERGER, E.—*Ber. dtsh. chem. Ges.* **33** : 1950, 3192 (1900).
- (11) LIMPRICHT, H.—*Ber. dtsh. chem. Ges.* **18** : 1404 (1885).
- (12) GRAEFF, F.—*Liebigs Ann.* **229** : 344 (1885).
- (13) SCHULTZ, G.—*Ber. dtsh. chem. Ges.* **17** : 464 (1884).
- (14) ELBS, K., and SCHWARTZ, B.—*J. prakt. Chem.* **63** : 567 (1901).
- (15) BAMBERGER, E.—*Ber. dtsh. chem. Ges.* **44** : 1967, footnote 3 (1911).
- (16) KEHRMANN, F., and BÜHLER, E.—*Ber. dtsh. chem. Ges.* **39** : 137 (1906).
- (17) v. AUWERS, K., BORSCHKE, E., and WELLER, R.—*Ber. dtsh. chem. Ges.* **54** : 1315 (1921).



- (18) McPHERSON, W., and BOORD, C.—*J. Amer. Chem. Soc.* **33** : 1531 (1911).
- (19) LIEBERMANN, C., and v. KOSTANECKI, S.—*Ber. dtsh. chem. Ges.* **17** : 876 (1884).
- (20) HEUMANN, K.—*Ber. dtsh. chem. Ges.* **5** : 913 (1872).
- (21) BAMBERGER, E.—*Ber. dtsh. chem. Ges.* **35** : 1620 (1902).
- (22) ANGELI, A.—*R.C. Accad. Lincei* (5) **22** : i, 356, 845 (1913).
- (23) BIGIAMI, D., and SABATELLI, V.—*Gazz. Chim. Ital.* **57** : 565 (1927).
- (24) WACKER, L.—*Liebigs Ann.* **317** : 376 (1901).
- (25) BREWSTER, C. M., and MILLAM, L. H.—*J. Amer. Chem. Soc.* **55** : 763 (1933).
- (26) ZININ, N.—*J. prakt. Chem.* **36** : 101 (1845).
- (27) KNIPSCHER, H. M.—*Rec. Trav. Chim. Pays-Bas* **22** : 15 (1903).
- (28) JACOBSON, P., HEBER, E., HENRICH, F., and SCHWARZ, C.—*Liebigs Ann.* **287** : 186 (1895).
- (29) DOWNES, A. M.—B.Sc. (Hons.) Thesis, University of Sydney (1946).
- (30) KRAUSE, M.—*Ber. dtsh. chem. Ges.* **32** : 124 (1899).
- (31) STOERMER, R., and WODARG, F.—*Ber. dtsh. chem. Ges.* **61** : 2330 (1928).
- (32) HEWITT, J. T., and MITCHELL, H. V.—*J. Chem. Soc.* **87** : 228 (1905).
- (33) WERNER, A., and STIASNY, E.—*Ber. dtsh. chem. Ges.* **32** : 3274 (1899).
- (34) BARSILOWSKY, J.—*Liebigs Ann.* **207** : 103 (1881).

# ALKALOIDS OF THE AUSTRALIAN RUTACEAE

## THE STRUCTURE AND REACTIONS OF ACRONYCIDINE

By F. N. LAHEY,\* J. A. LAMBERTON,† and J. R. PRICE‡

[Manuscript received November 7, 1949]

### Summary

Acronycidine,  $C_{15}H_{15}O_5N$ , is shown to be 4,5,7,8-tetramethoxyfurano-(2', 3'-2, 3)-quinoline (VI). It has been degraded to 2,4,5-trimethoxybenzoic acid by a series of reactions, one of which involves the conversion of 3-nitroso-5,7,8-trimethoxy-2,4-dihydroxyquinoline to 4,6,7-trimethoxyisatin by heating with mineral acid. This process appears to be of general application for the conversion of 2,4-dihydroxyquinolines to isatins. Acronycidine readily undergoes oxidative demethylation with either nitric or nitrous acid giving a 1,4-quinone. Quinones are likewise formed from several acronycidine derivatives.

### I. THE STRUCTURE OF ACRONYCIDINE

Acronycidine,  $C_{15}H_{15}O_5N$ , has been isolated from the bark of both *Acronychia baueri*(1) and *Melicope fareana*(2). It is a colourless, weak base containing four methoxyl groups but no methylimino group, is insoluble in alkali, cannot be acetylated and does not react with carbonyl reagents. It is hydrogenated to dihydroacronycidine,  $C_{15}H_{17}O_5N$ , and is oxidized by permanganate to an aldehyde,  $C_{14}H_{15}O_6N$  (VIII), and the corresponding acid,  $C_{14}H_{15}O_7N$  (IX). On heating with methyl iodide it is isomerized to isoacronycidine (X) which contains three methoxyl groups and one methylimino group. As these reactions are characteristic of the furanoquinoline alkaloids, dictamnine(3),  $\gamma$ -fagarine(4), and skimmianine(5), which have been found only in the Rutaceae, acronycidine would appear to be a tetramethoxyfuranoquinoline, a structure in accord with the analytical data.

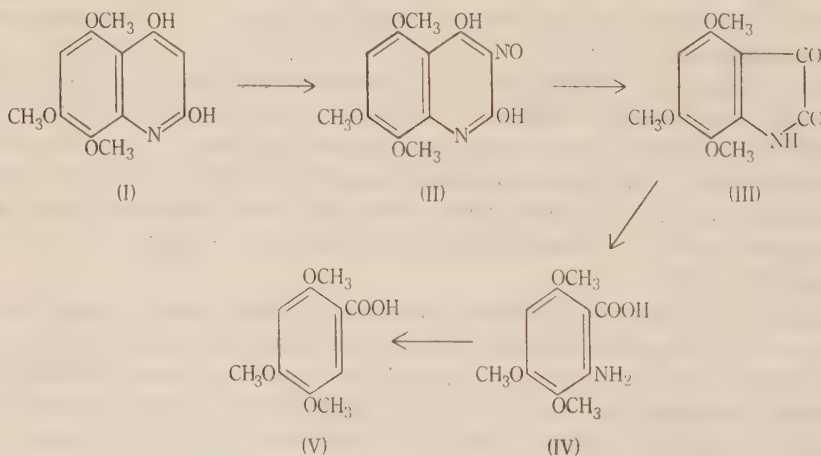
Oxidation of the known furanoquinoline alkaloids leads to acids which on heating with hydrochloric acid undergo simultaneous decarboxylation and demethylation of one methoxyl group, the products being 2,4-dihydroxyquinolines. Similarly the acid (IX) is decarboxylated and one methoxyl group demethylated by hydrochloric acid to a substance,  $C_{12}H_{13}O_5N$ , which is presumably a trimethoxy-2,4-dihydroxyquinoline. The orientation of the methoxyl groups is either 5,6,8- or 5,7,8- since on oxidative demethylation acronycidine gives a 1,4-quinone (*vide infra*).

\* Chemistry Department, University of Queensland.

† Chemistry Department, University of Melbourne.

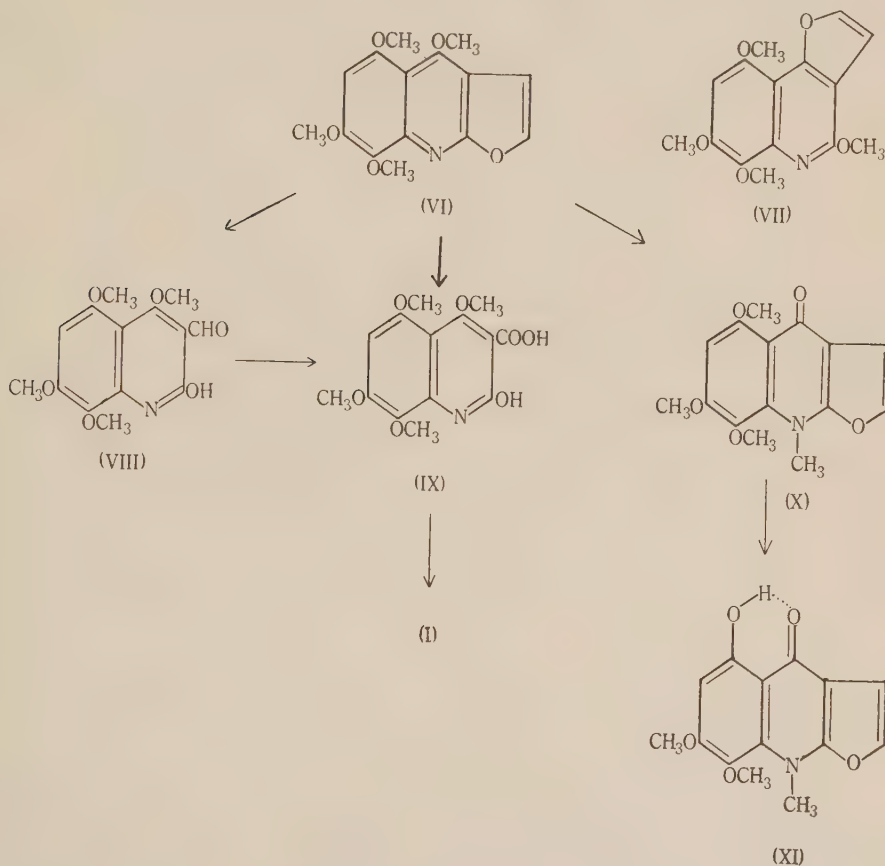
‡ Division of Industrial Chemistry, C.S.I.R.O., Melbourne.

Asahina and Nakanishi(6) and Berinzaghi *et al.*(7) have used Bischoff's reductive cyclization of an ethyl *o*-nitrobenzoylmalonate to synthesize the 2,4-dihydroxyquinolines derived from skimmianine and  $\gamma$ -fagarine. We have not been able to extend this to either of the required trimethoxy compounds because of difficulty in preparing the necessary trimethoxynitrobenzoic acids. Further, in model experiments based on the procedures of Meyer and Heimann(8) and Kämmerer(9) attempted syntheses of 2,4-dihydroxyquinolines by cyclization of ethyl malonanilates gave only dianilides, as has also been reported by Vaughan(10). Consequently an attempt was made to oxidize the presumed dihydroxyquinoline directly to the corresponding anthranilic acid. As this proved unsuccessful, the nitroso derivative of the dihydroxyquinoline was treated with aqueous sulphuric acid with the object of hydrolysing it to a triketotetrahydroquinoline which should oxidize more easily. The expected product was not obtained, but instead a substance,  $C_{11}H_{11}O_5N$ , which still contained the three methoxyl groups. This substance is insoluble in cold aqueous sodium bicarbonate, but dissolves in caustic alkalis giving a deep red solution, fading rapidly to yellow, from which it is regenerated on acidification. It reacts with carbonyl reagents, condenses with *o*-phenylenediamine to give a *phenazine*,  $C_{17}H_{15}O_3N_3$ , and is oxidized by alkaline hydrogen peroxide to a trimethoxyamino acid,  $C_{10}O_{13}O_5N$ , which is deaminated smoothly to 2,4,5-trimethoxybenzoic acid (asaronic acid, V). It follows that the substance  $C_{11}H_{11}O_5N$  is 4,6,7-trimethoxyisatin (III) and that the amino acid is 2-amino-3,4,6-trimethoxybenzoic acid (IV).



The reactions described above lead to the conclusion that acronycidine is either VI or VII. If VI is correct then the aldehyde and acid formed by oxidation of acronycidine have structures VIII and IX respectively, whereas VII would give rise to the analogous 2-methoxy-4-hydroxy aldehyde and acid. When the hydrochloride of *isoacronycidine* (X) is heated at its melting point, one methoxyl group is demethylated giving a compound,  $C_{14}H_{13}O_5N$  (XI, *norisoacronycidine*), which is insoluble in caustic soda, but forms a *monoacetyl* derivative

and can be remethylated to *isoacronycidine*. It is clearly comparable with the *nor*-compounds formed in the same manner from acronycine(11) and the *Melicope* alkaloids(12). Hence the pyridone carbonyl group of *isoacronycidine* must be situated *peri* to a methoxyl group in the benzenoid ring, thus eliminating structure VII for acronycidine. The linear tricyclic arrangement as in VI, has been established by Asahina, Ohta, and Inubuse(3) for dictamnine, but the accepted structures for  $\gamma$ -fagarine and skimmianine are based only on analogy.



The ultraviolet absorption spectra of skimmianine, acronycidine, *iso*-acronycidine, and *noriso*acronycidine have been measured. The absorption curves are illustrated in Figure 1 and the  $\lambda_{max}$  values recorded in Table 1.

The acronycidine absorption bears a general resemblance to that of skimmianine but the long-wave maximum has undergone a considerable displacement to longer wavelengths. Conversion to the *isocompound* has little effect on the absorption, but the long-wave maximum is displaced to longer wavelengths on formation of *noriso*acronycidine, that is on demethylation of the methoxyl group situated *peri* to the carbonyl. This is in accord with the findings of Brown and Lahey(13) in studies of the acridone alkaloids.



TABLE I

Substance	$\lambda_{max}$ (Å)	Log $\epsilon_{max}$	$\lambda_{max}$ (Å)	Log $\epsilon_{max}$
Skimmianine ..	3260	3.93	2480	4.91
Acronycidine ..	3460	3.96	2560	4.88
<i>iso</i> Acronycidine ..	3470	4.17	2530	4.69
<i>noriso</i> Acronycidine	3560	4.05	2530	4.69

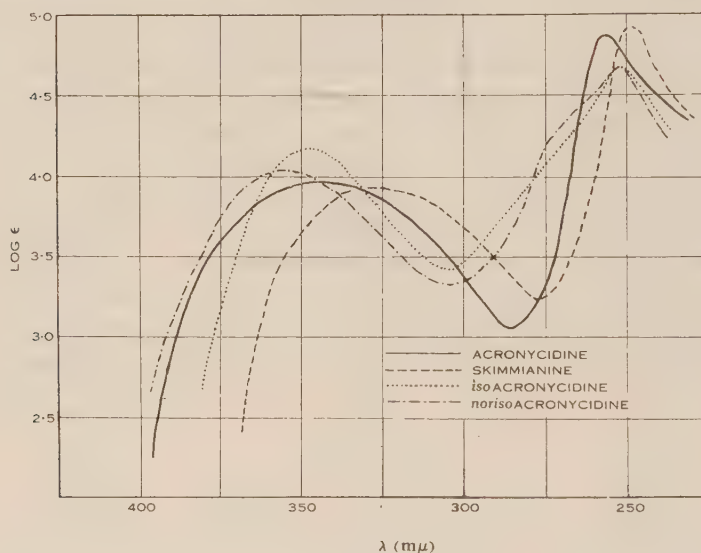


Fig. 1.

## II. THE CONVERSION OF 3-NITROSO-2,4-DIHYDROXYQUINOLINES TO ISATINS

Confirmation that the compound,  $C_{12}H_{13}O_5N$ , is correctly formulated as a trimethoxy-2,4-dihydroxyquinoline (I) is found in the preparation of isatin itself from 3-nitroso-2,4-dihydroxyquinoline under the same conditions and likewise by the preparation from skimmianine of 6,7-dimethoxyisatin, identified by comparison with an authentic specimen prepared by the method of Gulland *et al.*(14). It was discovered subsequently that the conversion of 3-nitroso-2,4-dihydroxyquinoline to isatin and hydroxylamine by boiling with concentrated hydrochloric acid has been reported by Baeyer and Homolka(15), and that Friedländer and Müller(16) similarly converted the N-methyl compound to N-methylisatin. Baeyer and Homolka, however, provided no experimental evidence for the statement that hydroxylamine is produced, and in the case of the trimethoxynitrosodihydroxyquinoline, we were unable to detect any hydroxylamine in the reaction mixture. Instead, ammonium salts were present. The yield of trimethoxyisatin is good and it is accompanied by the evolution

of an approximately equivalent amount of carbon dioxide, and a trace of carbon monoxide. Apart from the carbon monoxide, these products would seem to be satisfactorily accounted for by a mechanism involving a Beckmann rearrangement, followed by hydrolysis of the resulting cyclic urea derivative to the isatinic acid. However, this interpretation does not adequately cover the case of 3-nitroso-2,4-dihydroxyquinoline itself. Here the yield of isatin is low, *c.* 25 per cent., and it is accompanied by an equal amount of isatin- $\beta$ -oxime, while the remainder of the starting material is not accounted for. Presumably the isatoxime is formed from isatin and hydroxylamine, the latter arising by hydrolysis of the original nitroso compound. Such hydrolysis would also give triketotetrahydroquinoline which might degrade further. It seems that with 2,4-dihydroxyquinoline two different reactions may be involved.

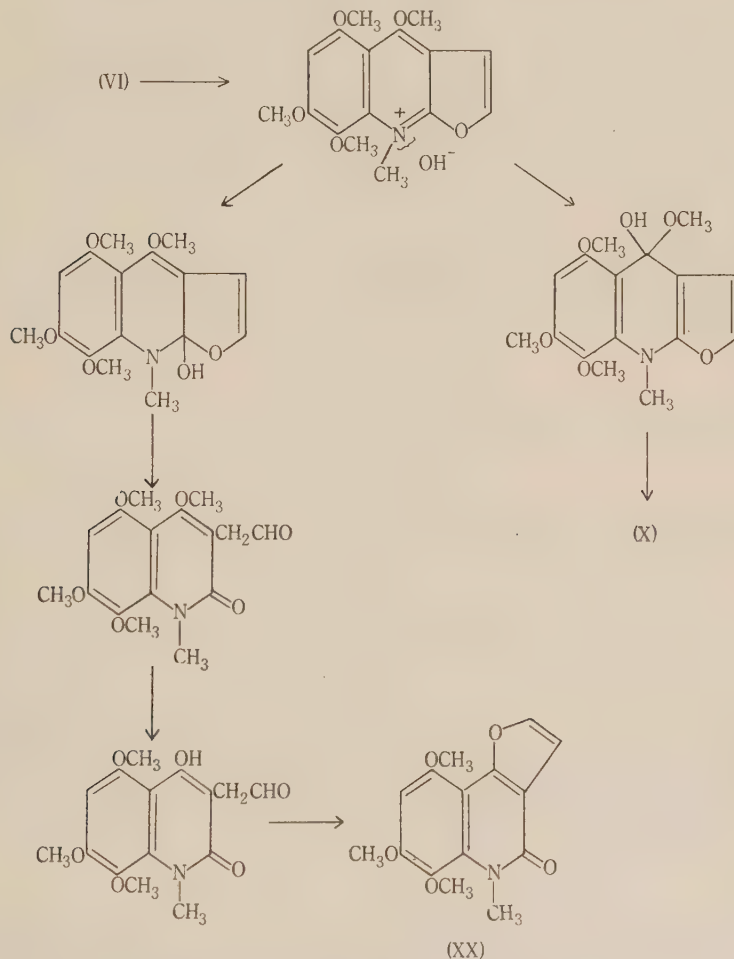
### III. THE REACTIONS OF ACRONYCIDINE

Berinzaghi *et al.* (17) have reported that the 4-methoxyl group in skimmianine and  $\gamma$ -fagarine undergoes an alkoxyl interchange on heating with alcoholic alkalis. Acronycidine behaves similarly with *isopropanolic* potash, giving a substance  $C_{17}H_{19}O_5N$ . It likewise reacts with ethanolic and *n*-propanolic potash, but the products could not be satisfactorily purified. In these reactions there is also formed the potassium salt of a phenol,  $C_{14}H_{13}O_5N$ , containing three methoxyl groups. The same phenol is obtained in good yield when acronycidine is heated with alcoholic hydrochloric acid. Oxidation of its *acetyl* derivative gives an acid which is converted to the trimethoxy-2,4-dihydroxyquinoline (I) by boiling mineral acid. Consequently it is the 4-methoxyl group which is demethylated by either alkali or acid and the phenol is 4-hydroxy-5,7,8-trimethoxyfurano-(2',3'-2,3)quinoline (XII). Methylation converts it to *iso*-acronycidine while ozonolysis in chloroform solution gives a *phenolic aldehyde*,  $C_{13}H_{13}O_6N$ , presumably XIII. Ozonolysis of acronycidine in chloroform solution leads only to the formation of an intractable gum, but under the same conditions *iso*acronycidine yields the *aldehyde*,  $C_{14}H_{15}O_6N$  (XVII). Neither this aldehyde nor the corresponding acid results from the oxidation of *iso*acronycidine with permanganate in acetone, the product is then the *hydroxy-N-methylquinolone* (XVIII).

The most characteristic property of acronycidine is the facility with which it undergoes oxidative demethylation to a yellow *quinone*,  $C_{13}H_9O_5N$  (XIV), when treated with either nitric or nitrous acids. This quinone is also formed when acronycidine is ozonized in acetic acid solution. It contains two methoxyl groups, gives a *diacetyldihydro* compound and since it does not react with *o*-phenylenediamine, is evidently a 1,4-quinone. Reduction is difficult under acid conditions, but is effected readily with sodium hydrosulphite (dithionite). However, the quinol has not been isolated on account of the ease with which it oxidizes in contact with air. The quinone does not dissolve in cold alkali, but does so on warming, giving a red solution from which on acidification is precipitated an orange *hydroxyquinone*,  $C_{12}H_7O_5N$ , which contains one methoxyl group, and condenses with *o*-phenylenediamine to give a violet *phenazine*. The structure XV follows from the method of preparation.



as the sole product in some experiments. In other experiments, however, a second product is formed which is isomeric with acronycidine and *isoacronycidine*, but more weakly basic than either. This isomer is best prepared, though in small yield,\* by treating acronycidine with dimethyl sulphate and alkali. It can only be *1-methyl-5,7,8-trimethoxyfurano-(3',2'-3,4)-quinol-2-one* (XX) as it gives on ozonolysis the same aldehyde (XVII) as *isoacronycidine*. The following mechanism is suggested to account for its formation. Unfortunately insufficient skimmianine was available to investigate the formation from it of a compound corresponding to XX.



#### IV. THE PHARMACOLOGY OF ACRONYCIDINE

The following report on the pharmacology of acronycidine has been furnished by Associate Professor F. H. Shaw, Department of Physiology, University of Melbourne. Test solutions were of the hydrochloride in saline or frog's Ringers

\* The main product is *isoacronycidine*.



solution. Acronycidine is relatively non-toxic in mice, the toxic dose lying above 150 mg./kg. The symptoms in mice indicated some loss of muscular coordination although righting reflexes appeared undisturbed. On an isolated frog's heart (Straub's heart preparation) a concentration of 1:1000 produced partial heart block finally becoming complete, with recovery on removal of the alkaloid solution. Concentrations of 1:50,000 produced immediate relaxation and abolished rhythmic contractions of an isolated rat intestine (Magnus's method). With isolated uterus (guinea pig) results were variable; sometimes there was no effect, sometimes contraction in concentrations ranging from 1:4000 to 1:20,000.

Acronycidine showed no curare action on isolated nerve muscle (toad) in concentrations of 1:5000. Higher concentrations produced a very marked slow contraction of the gastrocnemius muscle. This effect on voluntary muscle—producing contraction in concentrations of 1:1000—has been confirmed by experiments on the rectus abdominus muscle of the toad. 30 mg. caused a very slight decrease in the blood pressure of a cat, and this was followed by immediate recovery.

In its two most marked properties, the effect on voluntary muscle and causing cessation of heart beat, acronycidine resembles skimmianine(18).

## V. EXPERIMENTAL

All melting points are corrected unless otherwise stated. Microanalyses were carried out by R. B. Bradbury, N. L. Lottkowitz and K. Tettweiler.

### (a) *Properties of Acronycidine*

Acronycidine is almost insoluble in water and light petroleum, sparingly soluble in benzene and ethyl acetate, and very soluble in chloroform. It is conveniently crystallized from methanol or ethanol, from which it separates as colourless needles or prisms, m.p. 136.5–137.5 °C.

Found:  $\text{CH}_3\text{N}$ , nil;  $\text{CH}_3\text{O}$ , 42.8%; mol. wt. (Rast), 297;  $[\alpha]_{\text{D}}^{18}$ , 0° in chloroform.

Calculated for  $\text{C}_{15}\text{H}_{15}\text{O}_5\text{N}$ :  $\text{CH}_3\text{O}$ , 42.9% (four methoxyls); mol. wt., 289.

Acronycidine is insoluble in aqueous alkali but soluble in dilute acids. The *hydrochloride* (m.p. 120–121 °C. with decomposition), which separates as pale yellow needles when hydrochloric acid is added to an acetone solution of the base, is unstable and slowly loses hydrogen chloride at room temperature. It dissolves in water, but acronycidine may be recovered from the solution by extraction with chloroform. Acronycidine is also extracted slowly from more strongly acid solutions. *Acronycidine picrate* separates from methanol as long yellow needles, m.p. 181.5–182.5 °C.

Found: N, 11.1;  $\text{CH}_3\text{O}$ , 24.5%.

Calculated for  $\text{C}_{15}\text{H}_{15}\text{O}_5\text{N} \cdot \text{C}_6\text{H}_3\text{O}_7\text{N}_3$ : N, 10.8;  $\text{CH}_3\text{O}$ , 23.9%.

Acronycidine was recovered unchanged (m.p. and mixed m.p.) after refluxing with acetic anhydride and pyridine for 1 hour and after treatment at room temperature with benzoyl chloride in pyridine. It was also recovered unchanged after attempted preparation of an oxime and semicarbazone. No tractable product resulted when acronycidine was ozonized in chloroform solution.

### (b) *isoAcronycidine (X)*

Acronycidine (0.5 g.) was heated in a sealed tube at 100 °C. for 4 hours with excess methyl iodide (0.5 ml.). After opening the tube, the methyl iodide was allowed to evaporate and the residue crystallized from water. It separated as long, colourless needles, m.p. 172–173 °C.

Found (after drying in a vacuum at 110 °C.): C, 62.5; H, 5.3; N, 4.9; CH<sub>3</sub>O, 31.8; CH<sub>3</sub>N, 9.3%.

Calculated for C<sub>15</sub>H<sub>15</sub>O<sub>5</sub>N: C, 62.3; H, 5.2; N, 4.8; CH<sub>3</sub>O, 32.2 (three methoxyls); CH<sub>3</sub>N, 10.0%.

isoAcronycidine is very soluble in alcohols, dilute solutions having a strong blue fluorescence. The *hydrochloride*, yellow needles, m.p. 188–190 °C. (uncorr.) with decomposition, prepared in the same manner as that of acronycidine, was decomposed by water.

#### (c) Dihydroacronycidine

Acronycidine was hydrogenated in ethanol solution over Raney nickel at atmospheric pressure. Repeated crystallization of the product from alcohol, and subsequently from ethyl acetate, gave *dihydroacronycidine* as thick white plates, m.p. 188.5–190.5 °C.

Found: C, 62.0; H, 6.0; N, 4.9; CH<sub>3</sub>O, 42.8%.

Calculated for C<sub>15</sub>H<sub>17</sub>O<sub>5</sub>N: C, 61.9; H, 5.8; N, 4.8; CH<sub>3</sub>O, 42.6%.

On heating with methyl iodide in a sealed tube, dihydroacronycidine, as would be expected, gave a mixture of products.

#### (d) Oxidation of Acronycidine

Acronycidine (4 g.) dissolved in acetone (200 ml.) was treated with powdered potassium permanganate (6 g.) at room temperature. When oxidation was complete, the mixture was filtered and the manganese dioxide precipitate extracted with boiling water. Acidification with acetic acid gave the *acid* (IX) which separated from alcohol as colourless needles, m.p. 210–212 °C. with decomposition when placed in the melting point bath at c. 200 °C. Yield 1.5 g.

Found: C, 53.8; H, 4.9; N, 4.4; CH<sub>3</sub>O, 39.7%.

Calculated for C<sub>14</sub>H<sub>15</sub>O<sub>7</sub>N: C, 54.4; H, 4.9; N, 4.5; CH<sub>3</sub>O, 40.1% (four methoxyls).

IX is soluble in aqueous sodium carbonate, but insoluble in bicarbonate. It is only sparingly soluble in most organic solvents.

The filtrate from the manganese compounds was evaporated to dryness and the residue, 0.4 g., crystallized from alcohol. VIII was obtained as fine cream needles, m.p. 219.5–220.5 °C.

Found: C, 57.0; H, 5.2; N, 4.8; CH<sub>3</sub>O, 41.8%.

Calculated for C<sub>14</sub>H<sub>15</sub>O<sub>6</sub>N: C, 57.3; H, 5.1; N, 4.8; CH<sub>3</sub>O, 42.3%.

VIII formed a 2,4-dinitrophenylhydrazone, dark red needles from acetic acid, m.p. 302–304 °C. (decomp., uncorr.).

Found: N, 15.0%.

Calculated for C<sub>20</sub>H<sub>19</sub>O<sub>9</sub>N<sub>5</sub>: N, 14.8%.

The aldehyde dissolved in acetone was oxidized by the gradual addition of an acetone solution of potassium permanganate. Extraction of the resulting manganese compounds with boiling water, acidification, and crystallization of the precipitated acid from alcohol gave IX, m.p. 210–212 °C. (decomp., immersed at c. 200 °C.). A mixture with an authentic specimen of IX behaved in the same way.

#### (e) Decarboxylation of IX

The acid IX (1 g.), suspended in hydrochloric acid (5 N; 40 ml.) was heated under reflux until completely dissolved (1½–2 hours). Long colourless needles, readily soluble in alcohol or water, separated on cooling. A qualitative test showed the presence of chlorine, the substance evidently being a *hydrochloride*. After crystallization from 10% hydrochloric acid, it melted at 197–198 °C. Yield 0.8–0.9 g.

Found: CH<sub>3</sub>O, 32.5%.

Calculated for C<sub>12</sub>H<sub>13</sub>O<sub>5</sub>N.HCl: CH<sub>3</sub>O, 32.3% (three methoxyls).

The free *base* was obtained by treating an aqueous solution of the hydrochloride with sodium

acetate, and extracting with chloroform. Evaporation of the chloroform solution and crystallization from ethyl acetate-light petroleum gave I as colourless needles, m.p. 231–232 °C.

Found: C, 57.0; H, 5.5; N, 5.5;  $\text{CH}_3\text{O}$ , 36.8%.

Calculated for  $\text{C}_{12}\text{H}_{13}\text{O}_5\text{N}$ : C, 57.4; H, 5.2; N, 5.6;  $\text{CH}_3\text{O}$ , 37.1% (three methoxyls).

Acetylation with acetic anhydride and pyridine gave a *monoacetyl* derivative, fine white needles from aqueous methanol, m.p. 204–206 °C.

Found: N, 4.7;  $\text{CH}_3\text{O}$ , 31.7%.

Calculated for  $\text{C}_{12}\text{H}_{12}\text{O}_4\text{N}(\text{OOC}\cdot\text{CH}_3)$ : N, 4.8;  $\text{CH}_3\text{O}$ , 31.7%.

Acidification of an ice-cold solution of I in 5% caustic soda containing sodium nitrite (1 mole) gave a red precipitate of the *nitroso* derivative (II). Purification was effected by repeated crystallization of the sodium salt from 10% caustic soda solution, followed by crystallization of the free nitroso compound from acetic acid. It was obtained as red needles, m.p. 241–243 °C. (decomp.).

Found: C, 51.7; H, 4.4; N, 9.8%.

Calculated for  $\text{C}_{12}\text{H}_{12}\text{O}_6\text{N}_2$ : C, 51.4; H, 4.3; N, 10.0%.

Oxidation of I with alkaline permanganate or with permanganate in the presence of magnesium sulphate gave only oxalic acid. It was oxidized by ammoniacal silver nitrate giving in good yield a yellow crystalline solid which appeared to be a mixture. This product dissolved in caustic soda giving a red solution which faded rapidly to yellow and thus probably contained some of the trimethoxyisatin (*vide infra*). It was not further investigated.

#### (f) 4,6,7-Trimethoxyisatin (III)

The nitroso compound (II; 2 g.) was suspended in sulphuric acid (30%; 80 ml.) and the mixture heated to boiling, when solution took place with vigorous frothing. Heating was continued for 15–20 minutes. On cooling dark red needles of the isatin separated. A further quantity was obtained by extraction of the filtrate with chloroform, and crystallization of the extracted material first from *n*-propanol, then from water. The trimethoxyisatin separated from water as yellow needles which darkened and decomposed above 260 °C. (uncorr.) and were completely molten at 270 °C. (uncorr.). Yield 0.9 g.

Found: C, 56.2; H, 4.8; N, 6.0;  $\text{CH}_3\text{O}$ , 38.8%.

Calculated for  $\text{C}_{11}\text{H}_{11}\text{O}_5\text{N}$ : C, 55.7; H, 4.6; N, 5.9;  $\text{CH}_3\text{O}$ , 39.2%.

The substance dissolved in caustic soda giving a deep red solution which faded rapidly to yellow. On acidification the isatin was regenerated (m.p. and mixed m.p.). It gave a *phenylhydrazone*, short bright yellow needles from alcohol, m.p. 220.5–221.5 °C.

Found: N, 13.1;  $\text{CH}_3\text{O}$ , 28.4%.

Calculated for  $\text{C}_{17}\text{H}_{17}\text{O}_4\text{N}_3$ : N, 12.8;  $\text{CH}_3\text{O}$ , 28.4%.

The *oxime*, pale yellow needles from water, melted at 210 °C. (decomp.).

Found: N, 11.0%.

Calculated for  $\text{C}_{11}\text{H}_{12}\text{O}_5\text{N}_2$ : N, 11.1%.

Treatment of the isatin in alcoholic solution with *o*-phenylenediamine gave a *phenazine*, microscopic yellow needles from alcohol, m.p. 281–282 °C. (decomp., uncorr.).

Found:  $\text{CH}_3\text{O}$ , 30.3%.

Calculated for  $\text{C}_{17}\text{H}_{15}\text{O}_3\text{N}_3$ :  $\text{CH}_3\text{O}$ , 30.1%.

#### (g) 2-Amino-3,4,6-Trimethoxybenzoic Acid (IV)

The trimethoxyisatin was oxidized according to the method of Mayer and Schulze(19), 0.5 g. being dissolved in a solution of sodium hydroxide (1.5 g.) in water (4.5 ml.) and heated on a water-bath. Hydrogen peroxide (30%) was added dropwise until a small portion of the solution no longer gave a red colour on acidification. After addition of crushed ice the solution



was carefully acidified with hydrochloric acid until the yellow precipitate which first formed began to redissolve. The separation of the amino acid was completed by addition of sodium acetate. The product crystallized from benzene-light petroleum (charcoal) as flat white needles, m.p. 127.5–128.5 °C. Yield 0.1 g.

Found: C, 52.8; H, 5.8; N, 6.4;  $\text{CH}_3\text{O}$ , 40.8%.

Calculated for  $\text{C}_{10}\text{H}_{13}\text{O}_5\text{N}$ : C, 52.9; H, 5.7; N, 6.2;  $\text{CH}_3\text{O}$ , 41.0%.

The substance gives a bright red precipitate after treatment with nitrous acid followed by alkaline  $\beta$ -naphthol. It is very soluble in hot water, but only sparingly soluble in cold.

(h) 2,4,5-Trimethoxybenzoic Acid (V)

Trimethoxyanthranilic acid (IV; 200 mg.) suspended in 10% hydrochloric acid (0.9 ml.) was treated at 0 °C. with sodium nitrite (62 mg.) in water (0.2 ml.). The resulting yellowish-brown solution was filtered into cold 30% hypophosphorous acid (6 ml.) and the mixture kept at 0 °C. After 3 hours no more gas was evolved and crystallization of the product was induced by scratching. 2,4,5-Trimethoxybenzoic acid separated from water as colourless needles, m.p. 143.5–144.5 °C., undepressed by admixture with an authentic specimen. Yield 145 mg.

Found: C, 56.8; H, 6.0;  $\text{CH}_3\text{O}$ , 43.7%.

Calculated for  $\text{C}_{10}\text{H}_{12}\text{O}_5$ : C, 56.6; H, 5.7;  $\text{CH}_3\text{O}$ , 43.9%.

The identity was confirmed by the preparation from the acid of 2,4,5-trimethoxynitrobenzene, m.p. and mixed m.p. 129–130 °C. (see Fabinyi and Széki 20). Authentic 2,4,5-trimethoxybenzoic acid was obtained by oxidation with alkaline permanganate of 2,4,5-trimethoxybenzaldehyde (21) prepared in 90% yield from 1,2,4-trimethoxybenzene by the procedure of Adams and Levine (22).

(i) norisoAcronycidine (XI)

The hydrochloride of isoacronycidine decomposed with evolution of gas when heated at its melting point. The residue, which solidified on cooling, was recrystallized from chloroform-alcohol and obtained as cream needles, m.p. 226–227 °C.

Found: C, 60.9; H, 4.8; N, 5.1;  $\text{CH}_3\text{O}$ , 22.6%.

Calculated for  $\text{C}_{14}\text{H}_{13}\text{O}_5\text{N}$ : C, 61.1; H, 4.7; N, 5.1;  $\text{CH}_3\text{O}$ , 22.5% (two methoxyls).

norisoAcronycidine is insoluble in caustic soda and is an extremely weak base. It is sparingly soluble in alcohol but more soluble in chloroform. Acetylation with acetic anhydride and pyridine gave the *acetate*, colourless prisms from xylene, m.p. 174–175 °C.

Found:  $\text{CH}_3\text{O}$ , 19.6;  $\text{CH}_3\text{CO}$ , 13.0%.

Calculated for  $\text{C}_{14}\text{H}_{12}\text{O}_4\text{N}(\text{OOC.CH}_3)$ :  $\text{CH}_3\text{O}$ , 19.6;  $\text{CH}_3\text{CO}$ , 13.6%.

Methylation of norisoacronycidine with dimethyl sulphate and caustic soda on the water-bath gave isoacronycidine, m.p. and mixed m.p. 172–173 °C.

(j) Action of Alcoholic Hydrochloric Acid on Acronycidine

Acronycidine (4 g.) dissolved in alcohol (50 ml.) containing hydrochloric acid (35%; 12 ml.) was heated under reflux for 6 hours. The crystalline solid, m.p. 145–146 °C. (uncorr. decomp.), which separated was a hydrochloride. It was dissolved in warm water and the *phenolic base* (XII) precipitated by addition of sodium acetate. XII crystallized from ethyl acetate as colourless prisms, m.p. 185.5–186.5 °C. Yield 2.8 g.

Found: C, 61.0; H, 4.9; N, 5.0;  $\text{CH}_3\text{O}$ , 33.8%.

Calculated for  $\text{C}_{14}\text{H}_{13}\text{O}_5\text{N}$ : C, 61.1; H, 4.7; N, 5.1;  $\text{CH}_3\text{O}$ , 33.8% (three methoxyls).

The substance dissolves in caustic soda, but the sodium salt is only sparingly soluble in the presence of excess alkali. In alkaline solution it couples readily with diazonium salts. Methylation with dimethyl sulphate and caustic soda gave isoacronycidine, m.p. and mixed m.p. 172–173 °C. XII formed a *benzoate*, long cream needles from aqueous alcohol which melted at 142.5–143.5 °C.



Found: C, 66.5; H, 4.8; N, 3.9%.

Calculated for  $C_{21}H_{17}O_8N$ : C, 66.5; H, 4.5; N, 3.7%.

Acetylation of XII gave the *acetate*, colourless needles from alcohol, m.p. 133.5–134.5 °C.

Found: C, 60.5; H, 4.8; N, 4.3%.

Calculated for  $C_{18}H_{15}O_8N$ : C, 60.6; H, 4.7; N, 4.4%.

This acetate was oxidized by potassium permanganate in acetone solution and the acetyl-free product, recovered from the manganese dioxide precipitate, crystallized from alcohol. It was obtained as colourless needles which melted with effervescence at temperatures up to 200 °C., depending on the rate of heating, then resolidified and again melted at 228–233 °C.

Found:  $CH_3O$ , 31.5%.

Calculated for  $C_{13}H_{13}O_7N$ :  $CH_3O$ , 31.5% (three methoxys).

On boiling with 5N hydrochloric acid the acid was converted to the hydrochloride of the trimethoxy-2,4-dihydroxyquinoline (I). It was identified by conversion to the nitroso compound (II), m.p. and mixed m.p. 241–243 °C. with decomposition. XII was ozonized in chloroform solution, the solution then concentrated under reduced pressure, and the ozonide decomposed by hot water. The tarry product was washed with methanol and the residue crystallized from alcohol giving the aldehyde (XIII), short colourless needles, m.p. 251–252 °C. (decomp.).

Found: C, 55.6; H, 4.8; N, 5.2;  $CH_3O$ , 33.3%.

Calculated for  $C_{13}H_{13}O_6N$ : C, 55.9; H, 4.7; N, 5.0;  $CH_3O$ , 33.3% (three methoxys).

XIII gave a *phenylhydrazone*, thin yellow needles from aqueous acetic acid, m.p. 241–242 °C.

Found: N, 11.4;  $CH_3O$ , 25.4%.

Calculated for  $C_{13}H_{13}O_5N_2$ : N, 11.4;  $CH_3O$ , 25.2% (three methoxys). XIII gave only oxalic acid when oxidized by potassium permanganate in the presence of magnesium sulphate.

#### (k) Action of isoPropanolic Potash on Acronycidine

Acronycidine (0.5 g.) was refluxed for 1 hour with 5% isopropanolic potash (20 ml.). The crystalline solid which separated on cooling was dissolved in water and the solution saturated with carbon dioxide. This precipitated the phenolic base (XII), m.p. and mixed m.p. 185–186 °C. Yield 0.1 g. after recrystallization from aqueous alcohol. The isopropanolic potash solution was poured into water and allowed to stand, when a crystalline solid separated. This was repeatedly crystallized from alcohol, giving colourless prisms, m.p. 114–115 °C. The mixed m.p. with acronycidine was 102–105 °C.

Found: C, 64.2; H, 6.2; N, 4.4%.

Calculated for  $C_{17}H_{19}O_5N$ : C, 64.4; H, 6.0; N, 4.4%.

Similar reactions carried out with ethanolic and *n*-propanolic potash both yielded the phenol (XII) (m.p. and mixed m.p.). In addition there resulted in each case a basic fraction whose properties indicated that ether interchange had taken place to some extent at least. Using ethanolic potash the product melted at 98–100 °C. and its picrate at 184.5–185.5 °C. The mixed m.p. of the base with acronycidine was ill-defined and intermediate between the two but that of the picrate with acronycidine picrate was depressed c. 10 °C.

#### (l) Oxidation of isoAcronycidine

(i) *With Potassium Permanganate.*—isoAcronycidine (3 g.) dissolved in acetone (350 ml.) was oxidized by the addition of powdered potassium permanganate (5.5 g.) and the reaction completed by warming on the water-bath. The precipitate was suspended in water and treated with sulphur dioxide giving a yellow solid which was filtered off and taken up in warm aqueous sodium carbonate. The filtered solution was extracted with chloroform, the extract evaporated to dryness and the residue repeatedly crystallized from water giving XVIII in poor yield as colourless needles, m.p. 190–191 °C.

Found: C, 58.8; H, 5.6; N, 5.2;  $CH_3O$ , 35.4%.

Calculated for  $C_{13}H_{15}O_5N$ : C, 58.9; H, 5.7; N, 5.3;  $CH_3O$ , 35.2% (three methoxyls).

(ii) *Ozonolysis*.—*iso*Acronycidine was ozonized in chloroform solution, the solvent removed under reduced pressure, and the ozonide decomposed by boiling water. The water-insoluble product was crystallized several times from methanol and XVII obtained as cream needles, m.p. 214–215 °C.

Found: C, 57.8; H, 5.3; N, 4.8;  $CH_3O$ , 31.9%.

Calculated for  $C_{14}H_{18}O_5N$ : C, 57.3; H, 5.1; N, 4.8;  $CH_3O$ , 31.7% (three methoxyls).

The substance gave a *phenylhydrazone*, fine yellow needles from aqueous acetic acid, m.p. 242–243 °C.

Found: C, 62.9; H, 5.5; N, 11.1%.

Calculated for  $C_{20}H_{21}O_5N_3$ : C, 62.7; H, 5.5; N, 11.0%.

#### (m) Preparation of Quinones

(i) *From Acronycidine*.—Acronycidine (5 g.) was suspended in water (20 ml.) and 68% nitric acid (20 ml.) added carefully. Solution of the alkaloid was followed by the separation of a yellow crystalline mass. After dilution with water the product was filtered. Crystallization from acetic acid or from water gave XIV as yellow needles which melted with decomposition at 299–300 °C. (uncorr.) when placed in a melting point bath at c. 280 °C.

Found: C, 60.0; H, 3.6; N, 5.5;  $CH_3O$ , 23.8%.

Calculated for  $C_{13}H_9O_5N$ : C, 60.2; H, 3.5; N, 5.4;  $CH_3O$ , 23.9% (two methoxyls).

The high purity of the initial product made recrystallization generally unnecessary; the yield was over 90%. If recrystallization were necessary, pyridine was the most convenient solvent. The preparation could be carried out equally well with 68% nitric acid, with warm 2% nitric acid or with nitrous acid. XIV was also obtained in small amount when ozonized oxygen was passed through a solution of acronycidine in acetic acid.

XIV was recovered unchanged after attempted acetylation, but underwent reductive acetylation on treatment with acetic anhydride, pyridine, and zinc dust. The resulting *diacetyl-dihydro* derivative crystallized from methanol as colourless needles, m.p. 234–235 °C. (uncorr.).

Found: C, 59.5; H, 4.5; N, 4.3;  $CH_3CO$ , 24.7%.

Calculated for  $C_{13}H_9O_3N(OOC.CH_3)_2$ : C, 59.1; H, 4.3; N, 4.1;  $CH_3CO$ , 24.9%.

XIV was not reduced by sulphur dioxide, but a colourless quinol, which reoxidized immediately in contact with air, was formed by treatment of an aqueous suspension with sodium hydrosulphite (dithionite). The quinone did not react with *o*-phenylenediamine in boiling alcoholic solution.

By alkaline hydrolysis of XIV the *hydroxyquinone* (XV) was produced. This was best effected by making XIV into a thin paste with water, heating with 5% caustic soda for 1–2 minutes, filtering rapidly, and acidifying the filtrate. If the time of contact with alkali was prolonged, decomposition took place. The crystalline solid which separated on acidification was recrystallized from acetic acid giving XV, dull yellow needles melting with decomposition at 300 °C. (uncorr.) when immersed in a bath at c. 295 °C. If placed in the bath at a lower temperature decomposition took place without melting.

Found: C, 58.4; H, 2.9; N, 5.8;  $CH_3O$ , 12.2%.

Calculated for  $C_{12}H_7O_5N$ : C, 58.8; H, 2.9; N, 5.7;  $CH_3O$ , 12.7% (one methoxyl).

The *acetyl* derivative prepared by gently warming XV with acetic anhydride and pyridine, was crystallized from acetic acid-ethyl acetate. It separated as glistening yellow plates which decomposed above 200 °C. without melting.

Found: N, 5.0;  $CH_3CO$ , 14.9%.

Calculated for  $C_{12}H_6O_4N(OOC.CH_3)$ : N, 4.9;  $CH_3CO$ , 15.0%.

The *triacetyldihydro* derivative, prepared by reductive acetylation, separated from methanol as colourless needles, m.p. 196–197 °C. (uncorr.).

Found: N, 3.8%.

Calculated for  $C_{12}H_6O_2N(OOC.CH_3)_3$ : N, 3.8%.

XV condensed with *o*-phenylenediamine in acetic acid solution to give a *phenazine*, dull orange needles from acetic acid which decomposed with partial melting above 230 °C.

Found: N, 13.4;  $CH_3O$ , 9.1%.

Calculated for  $C_{18}H_{11}O_3N_3$ : N, 13.2;  $CH_3O$ , 9.8%.

XV was soluble in sodium bicarbonate solution and was reduced by sodium hydrosulphite (dithionite) to a quinol which reoxidized in contact with air.

(ii) *From isoAcronycidine*.—*isoAcronycidine* was suspended in 5% nitric acid and warmed gently. The base dissolved and almost at once a brown solid\* separated. This was filtered off, the filtrate poured on to excess sodium bicarbonate and the bright red solution repeatedly extracted with chloroform. Evaporation of the chloroform and crystallization of the residue (with charcoal) from methanol gave fine reddish-orange needles of XIX, m.p. 250–251 °C. (decomp.).

Found: C, 60.1; H, 3.8; N, 5.4;  $CH_3O$ , 12.4%.

Calculated for  $C_{13}H_9O_5N$ : N, 60.2; H, 3.5; N, 5.4;  $CH_3O$ , 12.0% (one methoxyl).

The same substance, identified by m.p. and mixed m.p., was formed by treating *norisoacronycidine* (XI) with 35% nitric acid. It was reduced by sodium hydrosulphite (dithionite) but the quinol reoxidized in contact with air.

(iii) *From the Phenol (XII)*.—The phenol (XII) was suspended in a little water and sufficient 68% nitric acid added to effect solution. Dilution with water resulted in the separation of an orange-red solid which was crystallized from acetic acid giving XVI as orange-red prisms, which decomposed without melting above 245 °C.

Found: C, 58.9; H, 3.1; N, 5.9;  $CH_3O$ , 12.5%.

Calculated for  $C_{12}H_7O_5N$ : C, 58.8; H, 2.9; N, 5.7;  $CH_3O$ , 12.7% (one methoxyl).

The substance dissolved readily in sodium bicarbonate solution and was reduced by sodium hydrosulphite (dithionite). It formed an *acetyl* derivative, pale yellow needles from acetic acid, which decomposed above 200 °C.

Found: C, 58.2; H, 3.3; N, 5.2%.

Calculated for  $C_{12}H_6O_4N(OOC.CH_3)_3$ : C, 58.5; H, 3.1; N, 4.9%.

XVI reacted slowly with *o*-phenylenediamine in boiling alcohol or acetic acid. The resulting *phenazine* separated from the reaction mixture as violet plates which could not be recrystallized because of their very low solubility. Purification was effected by treatment with boiling acetic acid. The product decomposed without melting at c. 245 °C.

Found: N, 13.4%;  $CH_3O$ , nil.

Calculated for  $C_{17}H_9O_3N_3$ : N, 13.9%.

The *phenazine* dissolved in aqueous caustic soda giving a violet solution from which a red sodium salt was precipitated by excess alkali.

(iv) *From Dihydroacronycidine*.—Treatment of dihydroacronycidine with nitric acid (35%) gave a *quinone*, yellow needles from acetic acid, which darkened and decomposed without melting over the range 245–250 °C.

\* This solid crystallized from chloroform-alcohol as yellow prisms, m.p. 249–250 °C. It appeared to be a *nitronorisoacronycidine* whose formation would resemble the preparation by Fabinyi and Széki(20) of the oxime of 4,5-dimethoxy-1,2-benzoquinone from 2,4,5-trimethoxybenzoic acid and nitrous acid.

Found: C, 52.0; H, 4.3; N, 8.7;  $CH_3O$ , 19.5%.

Calculated for  $C_{14}H_{12}O_7N_2$ : C, 52.5; H, 3.8; N, 8.8;  $CH_3O$ , 19.4% (two methoxyls).



Found: C, 59.9; H, 4.5; N, 5.4;  $\text{CH}_3\text{O}$ , 23.4%.

Calculated for  $\text{C}_{13}\text{H}_{11}\text{O}_5\text{N}$ : C, 59.8; H, 4.2; N, 5.4;  $\text{CH}_3\text{O}$ , 23.8% (two methoxyls).

(n) *Action of Dimethyl Sulphate on Acronycidine*

Treatment of acronycidine with a large excess of dimethyl sulphate and caustic soda (25%) gave a viscous oil which crystallized on cooling. The reaction mixture was extracted with chloroform, the extract evaporated and the residue dissolved in hot methanol. On cooling, fine white needles separated. The substance XX melted at 219–220 °C. after recrystallization from methanol.

Found: C, 62.3; H, 5.1; N, 4.7;  $\text{CH}_3\text{O}$ , 32.5;  $\text{CH}_3\text{N}$ , 9.0%.

Calculated for  $\text{C}_{15}\text{H}_{15}\text{O}_5\text{N}$ : C, 62.3; H, 5.2; N, 4.8;  $\text{CH}_3\text{O}$ , 32.2 (for three methoxyls);  $\text{CH}_3\text{N}$ , 10.0%.

Further small amounts of XX separated on standing, but the yield was low and variable, 5–25%. The methanol filtrate on evaporation, and crystallization of the residue from water, gave *isoacronycidine*, m.p. and mixed m.p. 172–173 °C., in good yield. XX was ozonized in chloroform solution, the ozonide decomposed by boiling with water and the product, after crystallization from alcohol, identified as the aldehyde XVII. The melting point (214–215 °C.) and that of the phenylhydrazone (242–243 °C.) were not depressed by admixture with samples obtained from the ozonolysis of *isoacronycidine*. XX is a very weak base—it is more soluble in hot 30% hydrochloric acid than in hot water, but the free base crystallizes on cooling.

(o) *Reductive Methylation of Quinones*

(i) The quinone (XIX), dissolved in warm water, was treated with a little sodium hydro-sulphite (dithionite) and the resulting quinol methylated by the addition of caustic soda (20%) and dimethyl sulphate. The reaction mixture was extracted with chloroform, the solvent evaporated and the residue crystallized from methanol giving *norisoacronycidine* (XI), m.p. and mixed m.p. 225–226 °C. Evaporation of the methanol mother liquors and crystallization of the residue from water gave *isoacronycidine*, m.p. and mixed m.p. 172–173 °C.

(ii) The quinone (XVI), reductively methylated as in (i), gave only *isoacronycidine*, m.p. and mixed m.p. 172–173 °C.

(iii) Reductive methylation of the quinones XIV and XV resulted in the same two products. Crystallization from methanol gave in each instance a small amount of the substance (XX) identified by its m.p. and mixed m.p. The methanol mother liquors were evaporated and the residues crystallized from water giving in each instance *isoacronycidine*, m.p. and mixed m.p. 172–173 °C.

(p) *Action of Acids on the Nitroso Derivatives of 2,4-Dihydroxyquinolines*

(i) *2,4-Dihydroxyquinoline*.—3-Nitroso-2,4-dihydroxyquinoline, treated with boiling 30% sulphuric acid as described for the preparation of III, dissolved with effervescence, but no solid separated on cooling. The reaction mixture was extracted three times with chloroform, then repeatedly with ether. The combined chloroform extracts were concentrated and the residue thrice crystallized from water giving orange-red needles, m.p. 203–204 °C., not depressed by admixture with an authentic specimen of isatin. The identity was confirmed by the preparation of isatin- $\beta$ -phenylhydrazone, m.p. and mixed m.p. 213.5–214 °C. A cleaner product was obtained from the chloroform extracts by solution in caustic soda (red solution  $\rightarrow$  yellow), acidification and re-extraction. The yield of isatin was c. 25%.

The combined ether extracts were concentrated and the residue crystallized several times from water. The product, yellow needles, m.p. 223–224 °C. (decomp.) was isatin- $\beta$ -oxime. Yield c. 22%.

Found: N, 17.1%.

Calculated for  $\text{C}_8\text{H}_6\text{O}_2\text{N}_2$ : N, 17.3%.



Treatment with boiling acetic anhydride gave the diacetyl derivative, pale yellow needles from acetic anhydride, m.p. 183–184 °C., alone or mixed with an authentic specimen. The melting point of this compound has been reported by Borsche and Sander(23) as 175 °C. Isatin- $\beta$ -oxime prepared by treating a hot aqueous solution of isatin with hydroxylamine hydrochloride and sodium acetate, melted at 201–203 °C. in agreement with the figure reported by Baeyer and Comstock(24). Repeated crystallization from water, however, caused the melting point to rise to 216 °C. as found by Borsche and Sander(23). These latter authors succeeded in further increasing the melting point to 221–222 °C. by preparation and hydrolysis of the diacetate. We have found that material melting at 223–224 °C. can be obtained either by heating a lower melting sample (m.p. 216 °C.) with 30% sulphuric acid for a few minutes, or directly by preparing the oxime from a solution of isatin in hot 30% sulphuric acid—it should be noted that, prepared under these conditions from equimolecular proportions of isatin and hydroxylamine, the yield of oxime was quantitative.

The formation of isatin and its oxime from 3-nitroso-2,4-dihydroxyquinoline by the action of 30% sulphuric acid was found to be accompanied by the liberation of a substantial amount of carbon dioxide and a trace of carbon monoxide, the latter detected by passing the evolved gases through a tube containing silica gel impregnated with palladous chloride. The acid liquors at the conclusion of the reaction contained ammonium sulphate as shown by evolution of ammonia on basification. The solution could not have contained hydroxylamine as it was shown (see above) that under the conditions employed isatin reacts quantitatively with hydroxylamine. The conversion of 3-nitroso-2,4-dihydroxyquinoline to isatin and isatin- $\beta$ -oxime could be brought about by more dilute acid, for instance, 10% hydrochloric acid gave substantially the same result as 30% sulphuric acid.

(ii) *7,8-Dimethoxy-2,4-Dihydroxyquinoline*.—3-Nitroso-7,8-dimethoxy-2,4-dihydroxyquinoline (prepared from the alkaloid skimmianine) was refluxed in 30% sulphuric acid for 1 hour. The cooled reaction mixture was extracted several times with chloroform, the combined chloroform extracts evaporated and the residue dissolved in caustic soda (red solution changing to yellow). The solution was acidified and extracted with chloroform, the extracts evaporated and the residue crystallized from water giving flat orange needles, m.p. 213–215 °C. These were identified as 6,7-dimethoxyisatin by comparison (mixed m.p., no depression) with a synthetic specimen(14).

Found:  $\text{CH}_3\text{O}$ , 29.7%.

Calculated for  $\text{C}_{10}\text{H}_9\text{O}_4\text{N}$ :  $\text{CH}_3\text{O}$ , 30.0% (two methoxyls).

The semicarbazone melted with decomposition at 253–255 °C. as reported by Gulland *et al.*(14).

(iii) *5,7,8-Trimethoxy-2,4-Dihydroxyquinoline*.—Conversion of the nitroso compound (II) to the trimethoxyisatin, as described above, was accompanied by the liberation of carbon monoxide and carbon dioxide. Only a trace of carbon monoxide was detected, but the amount of carbon dioxide was approximately one mole per mole pure trimethoxyisatin recovered. No oxime, as in the case of 3-nitroso-2,4-dihydroxyquinoline, could be detected. Basification of the acid solution after extraction of the trimethoxyisatin-liberated ammonia, but hydroxylamine was absent as shown by the non-formation of isatin- $\beta$ -oxime when another portion of the acid solution was treated with isatin.

The result of the reaction was unaffected by carrying it out in an atmosphere of nitrogen. It also proceeded smoothly with 10% hydrochloric acid in place of 30% sulphuric acid, but was extremely slow with 3% hydrochloric acid.

#### (q) Ultraviolet Absorption Spectra

The ultraviolet absorption spectra were measured in absolute alcoholic solution by means of a Spekker ultraviolet spectrophotometer in conjunction with a Hilger medium quartz spectrograph.

## VI. ACKNOWLEDGMENTS

The work described in this paper was carried out as part of the research programme of the Division of Industrial Chemistry, C.S.I.R.O., conducted in collaboration with the Chemistry Department, University of Melbourne.

The authors are indebted to Mr. L. J. Webb, Division of Plant Industry, C.S.I.R.O., for obtaining supplies of the necessary plant materials, to Professor E. J. Hartung for providing accommodation and facilities in the Chemistry Department, University of Melbourne, and to Associate Professor F. H. Shaw for reporting on the pharmacology of acronycidine.

## VII. REFERENCES

- (1) LAHEY, F. N., and THOMAS, W. C.—*Aust. J. Sci. Res. A* **2**: 423 (1949).
- (2) PRICE, J. R.—*Aust. J. Sci. Res. A* **2**: 249 (1949).
- (3) ASAHINA, Y., OHTA, T., and INUBUSE, M.—*Ber. dtsh. chem. Ges.* **63**: 2045 (1930).
- (4) DEULOFEU, V., LABRIOLA, R., and DE LANGHE, J.—*J. Amer. Chem. Soc.* **64**: 2326 (1942).
- (5) ASAHINA, Y., and INUBUSE, M.—*Ber. dtsh. chem. Ges.* **63**: 2052 (1930).
- (6) ASAHINA, Y., and NAKANISHI, S.—*Ber. dtsh. chem. Ges.* **63**: 2057 (1930).
- (7) BERINZAGHI, B., MURUZABAL, A., LABRIOLA, R., and DEULOFEU, V.—*J. Org. Chem.* **10**: 181 (1945).
- (8) MEYER, A., and HEIMANN, P.—*Compt. Rend.* **204**: 1204 (1937).
- (9) KÄMMERER, H.—German Patent 505,798; *Chem. Zentr.* **102**: I, 2679 (1931).
- (10) VAUGHAN, W. R.—*J. Amer. Chem. Soc.* **68**: 324 (1946).
- (11) BROWN, R. D., DRUMMOND, L. J., LAHEY, F. N., and THOMAS, W. C.—*Aust. J. Sci. Res. A* **2**: 627 (1949).
- (12) CROW, W. D., and PRICE, J. R.—*Aust. J. Sci. Res. A* **2**: 255 (1949).
- (13) BROWN, R. D., and LAHEY, F. N.—*Aust. J. Sci. Res. A* **3** (in press).
- (14) GULLAND, J. M., ROBINSON, R., SCOTT, J., and THORNLEY, S.—*J. Chem. Soc.* **1929**: 2932 (1929).
- (15) BAEYER, A., and HOMOLKA, B.—*Ber. dtsh. chem. Ges.* **16**: 2216 (1883).
- (16) FRIEDLÄNDER, P., and MÜLLER, F.—*Ber. dtsh. chem. Ges.* **20**: 2009 (1887).
- (17) BERINZAGHI, B., DEULOFEU, V., LABRIOLA, R., and MURUZABAL, A.—*J. Amer. Chem. Soc.* **65**: 1357 (1943).
- (18) HONDA, J.—*Arch. Exp. Path. Pharm.* **52**: 83 (1904).
- (19) MAYER, F., and SCHULZE, R.—*Ber. dtsh. chem. Ges.* **58**: 1465 (1925).
- (20) FABINYI, R., and SZÉKI, T.—*Ber. dtsh. chem. Ges.* **39**: 3679 (1906).
- (21) GATTERMANN, L., and EGGERS, F.—*Ber. dtsh. chem. Ges.* **32**: 289 (1899).
- (22) ADAMS, R., and LEVINE, I.—*J. Amer. Chem. Soc.* **45**: 2373 (1923).
- (23) BORSCHKE, W., and SANDER, W.—*Ber. dtsh. chem. Ges.* **47**: 2815 (1914).
- (24) BAEYER, A., and COMSTOCK, W.—*Ber. dtsh. chem. Ges.* **16**: 1706 (1883).

## CORRIGENDA

## VOLUME 1, NUMBER 1

Page 45, line 8 from bottom: *For* descending *read* ascending.

Page 45, line 7 from bottom: *For* ascending *read* descending.

## VOLUME 2, NUMBER 1

Page 98, 3 lines from bottom: *For*  $\alpha = \frac{1}{1 + \frac{(H^+)}{k_{OH}} + \frac{(H^+)}{k_{OH}} + \frac{(H^+)}{k_{NH}}}$

*read*  $\alpha = \frac{1}{1 + \frac{(H^+)}{k_{OH}} + \frac{(H^+)}{k_{OH}} \cdot \frac{(H^+)}{k_{NH}}}$

Page 127, line 8 from bottom: *For*  $C_{20}H_{20}O_4$  *read*  $C_{20}H_{24}O_4$ .

Page 129, line 4: *For*  $C_{20}H_{20}O_4$  *read*  $C_{20}H_{24}O_4$ .

# ESTIMATION OF THE ERROR AT A POINT ON A LEAST-SQUARES CURVE

By P. G. GUEST\*

[*Manuscript received December 13, 1949*]

## *Summary*

The estimation of the errors in the values obtained by fitting a least-squares polynomial to a number of equally-spaced observations of a variable quantity is discussed. Curves and tables of the weight function are obtained.

## I. INTRODUCTION

The problem of finding the curve of best fit for a series of equally-spaced observations has been thoroughly investigated by many writers: the estimation of the standard error or the "probable" error at a point on such a curve of best fit has been discussed by Schultz(1) and by Birge(2).

Schultz obtains formulae for the standard error at any point, and for the positions of the maxima and minima of the standard error, in terms of the Gauss-Doolittle coefficients  $[\alpha\alpha]$  etc. By fitting polynomials of degree 1, 2, and 3 to a set of observations, he obtains curves showing the function and its standard error for a specific problem.

The formulae for the errors are obtained much more simply in terms of orthogonal polynomials. Birge gives a Model Form by means of which the probable errors for any given problem may be obtained.

The purpose of the present paper is to discuss the manner in which the estimate of the error of the fitted value depends on the location of the point on the curve. It is shown that the curves giving the variation in the standard error or the "probable" error have practically the same form for all values of  $n$ , the number of observations. The locations of the maxima and minima of the standard error are determined. It is shown that the "weight function" can be put in a form which is largely independent of  $n$ . Tables are constructed from which the weight function can be read off directly.

## II. ORTHOGONAL REPRESENTATION

We consider the fitting of a polynomial of degree  $p$  to  $n$  equally-spaced observations. Following the notation of Birge(2), we denote the observations by  $y(\varepsilon)$ , where  $\varepsilon$  takes the  $n$  integral or half-integral values  $-(n-1)/2$  to  $+(n-1)/2$ . The curve of best fit for these observations is then most conveniently expressed in the form

$$u_p(\varepsilon) = \sum_{j=0}^p a_j T_j(\varepsilon),$$

\* Physics Department, University of Sydney.



where  $T_j(\varepsilon)$  is the orthogonal Tchebycheff polynomial of degree  $j$  in  $\varepsilon$ , and

$$a_j = \sum_{\varepsilon} y(\varepsilon) T_j(\varepsilon) / \sum_{\varepsilon} T_j^2(\varepsilon).$$

More usually, polynomials  $V_j(\varepsilon)$ —in Fisher's notation,  $\xi'_j(\varepsilon)$ —are employed. These are multiples of the  $T_j(\varepsilon)$  by a suitable least common factor  $\lambda_j$  such that the values  $V_j(\varepsilon)$  are all integers. Values of  $V_j(\varepsilon)$  are tabulated by Birge for values of  $j$  to 5,  $n$  to 30; by Fisher and Yates(3) for  $j$  to 5,  $n$  to 75; by Anderson and Housman(4) for  $j$  to 5,  $n$  to 104; and by van der Reyden(5) for  $j$  to 9,  $n$  to 52.

The estimate of the standard error of the fitted value  $u_p(\varepsilon)$  is given by

$$\begin{aligned} s[u_p(\varepsilon)] &= \left\{ \sum_{j=0}^p T_j^2(\varepsilon) / M_j \right\}^{\frac{1}{2}} s_p \\ &= \left\{ \sum_{j=0}^p V_j^2(\varepsilon) / N_j \right\}^{\frac{1}{2}} s_p, \end{aligned}$$

where

$$\begin{aligned} M_j &= \sum_{\varepsilon} T_j^2(\varepsilon) \\ N_j &= \sum_{\varepsilon} V_j^2(\varepsilon) \end{aligned}$$

and  $s_p$  is the estimate of the standard error of an observation defined in the usual way by the equation

$$s_p = \left\{ \sum_{\varepsilon} v_p^2(\varepsilon) / (n - p - 1) \right\}^{\frac{1}{2}},$$

the quantities  $v_p$  being the residuals.

In the physical sciences it is customary(7) to employ the concept of probable error  $r_p$ , defined by the equation

$$r_p = 0.6745 s_p.$$

Since this paper is addressed primarily to physicists, we shall for the remainder of the paper use the probable error concept. We shall be concerned with the ratio of the error at a point on the curve to the error of an observation, and the discussion will be unaltered whether standard errors or probable errors are used.

The probable error of the fitted value  $u_p(\varepsilon)$  is then

$$\begin{aligned} r[u_p(\varepsilon)] &= \left\{ \sum_{j=0}^p T_j^2(\varepsilon) / M_j \right\}^{\frac{1}{2}} r_p \\ &= \left\{ \sum_{j=0}^p V_j^2(\varepsilon) / N_j \right\}^{\frac{1}{2}} r_p, \end{aligned}$$

where  $r_p$  is the probable error of an observation.

Using the tabulated values of  $V_j(\varepsilon)$  and  $N_j$ , Birge(2) gives a Model Form by means of which the probable errors of the fitted values may be obtained. The process consists in the formation of the separate terms  $V_j^2(\varepsilon)/N_j$  and their summation.

### III. PROBABLE ERROR FORMULAE

The polynomials  $T_j(\varepsilon)$  and the sums  $M_j$  of their squared values may be written as explicit functions of the variable  $\varepsilon$ . The appropriate formulae are

given by Birge(2) and by Allen(6). Changing the variable to  $k=2\varepsilon/n$ , and taking a factor  $n^{-\frac{1}{2}}$  outside the bracket, the probable error can be put in the form

$$\begin{aligned} r[u_p(\varepsilon)] &= \frac{1}{\sqrt{n}} \left\{ \sum_{j=0}^p \tau_{j0}(k, n) \right\}^{\frac{1}{2}} r_p \\ &= \frac{1}{\sqrt{n}} \rho_{p0}(k, n) \cdot r_p, \text{ say.}^* \end{aligned}$$

The quantities  $\tau_{j0}(k, n)$  are given by the expressions

$$\begin{aligned} \tau_{00}(k, n) &= 1 \\ \tau_{10}(k, n) &= 3k^2 D_1 \\ \tau_{20}(k, n) &= \frac{5}{4} \{3k^2 - (1 - v^2)\}^2 \cdot D_2 \\ \tau_{30}(k, n) &= \frac{7}{4} k^2 \{5k^2 - (3 - 7v^2)\}^2 \cdot D_3 \\ \tau_{40}(k, n) &= \frac{9}{64} \{35k^4 - (30 - 130v^2)k^2 + 3(1 - v^2)(1 - 9v^2)\}^2 \cdot D_4 \\ \tau_{50}(k, n) &= \frac{11}{64} k^2 \{63k^4 - 70(1 - 7v^2)k^2 + (15 - 230v^2 + 407v^4)\}^2 \cdot D_5, \end{aligned}$$

where

$$v^2 = 1/n^2$$

$$D_j = 1/(1 - v^2)(1 - 4v^2) \dots (1 - j^2v^2).$$

It is seen that the quantities  $\tau(k, n)$ , and hence the quantities  $\rho(k, n)$ , vary only slowly with  $n$ ; for large  $n$  powers of  $v$  may be neglected.

The values of the variable  $\varepsilon$  may be divided into two groups. The region of interpolation comprises the values  $|\varepsilon| < \frac{1}{2}n$ , i.e.  $|k| < 1$ . The region of extrapolation comprises the values  $|\varepsilon| > \frac{1}{2}n$ , i.e.  $|k| > 1$ .

In the region of interpolation the variation in  $\rho_{p0}(k, n)$  is comparatively small. Table 1 gives the values  $\rho_{p0}(k, n)$  for  $|k| \leq 1$  and various selected values of  $n$ . Intermediate values may be obtained by linear interpolation between the tabulated values. The error arising from interpolation is generally less than 1 per cent., and never exceeds 2 per cent.

In the region of extrapolation  $\rho_{p0}(k, n)$  may be split up into two parts. We write

$$\begin{aligned} \rho_{p0}(k, n) &= \rho_{p0}(k) \cdot \varphi_p(n), \\ \rho_{p0}(k) &= \rho_{p0}(k, \infty), \\ \varphi_p(n) &= D_p^{\frac{1}{2}}. \end{aligned}$$

where

These functions are given in Table 2. The error introduced by splitting  $\rho_{p0}(k, n)$  into two factors is less than 1 per cent. for  $n \geq 12$ . When  $n < 12$ , the error is somewhat greater for values of  $k$  near unity; it is always less than 5 per cent. except in the single case  $n=7$ ,  $p=5$ .

\* More generally, the probable error of the  $q$ th order differential coefficient at the point  $k$  can be expanded as the square root of sums of terms  $\tau_{jq}(k, n)$ . The zero order coefficient is just the fitted value. Since it is intended to use the expressions  $\tau_{jq}(k, n)$  and  $\rho_{jq}(k, n)$  in a future paper, it is convenient to use here the forms  $\tau_{j0}(k, n)$  and  $\rho_{j0}(k, n)$ , the subscript 0 simply indicating that the expressions refer to the fitted value.

TABLE 1  
VALUES OF  $\rho_{p,0}(k,n)$

$n$ $k$	$p=1$ $\infty-7$	$\infty$	14	$p=2$ 11	8	7
0.00	1.00	1.50	1.51	1.51	1.52	1.53
0.05	1.00	1.50	1.50	1.51	1.52	1.52
0.10	1.01	1.49	1.49	1.50	1.50	1.51
0.15	1.03	1.47	1.47	1.48	1.49	1.49
0.20	1.06	1.44	1.45	1.45	1.46	1.47
0.25	1.09	1.42	1.42	1.43	1.43	1.44
0.30	1.13	1.39	1.39	1.40	1.40	1.41
0.35	1.17	1.37	1.37	1.37	1.37	1.38
0.40	1.22	1.35	1.35	1.35	1.35	1.36
0.45	1.27	1.34	1.34	1.34	1.35	1.35
0.50	1.32	1.35	1.35	1.35	1.36	1.36
0.55	1.38	1.38	1.39	1.39	1.39	1.39
0.60	1.44	1.44	1.45	1.45	1.45	1.45
0.65	1.51	1.54	1.54	1.54	1.55	1.55
0.70	1.57	1.66	1.66	1.67	1.68	1.68
0.75	1.64	1.81	1.82	1.83	1.84	1.85
0.80	1.71	1.99	2.01	2.01	2.03	2.05
0.85	1.78	2.21	2.22	2.23	2.26	2.27
0.90	1.85	2.45	2.47	2.48	2.51	2.53
0.95	1.93	2.71	2.74	2.75	2.79	2.81
1.00	2.00	3.00	3.03	3.05	3.09	3.12

$n$ $k$	$\infty$	25	18	$p=3$ 14	11	8	7
0.00	1.50	1.50	1.50	1.51	1.51	1.52	1.53
0.05	1.51	1.51	1.51	1.52	1.52	1.53	1.54
0.10	1.54	1.54	1.54	1.54	1.55	1.56	1.57
0.15	1.58	1.58	1.58	1.59	1.59	1.61	1.62
0.20	1.62	1.63	1.63	1.64	1.64	1.67	1.68
0.25	1.67	1.68	1.68	1.69	1.70	1.72	1.74
0.30	1.72	1.73	1.73	1.74	1.75	1.77	1.79
0.35	1.76	1.76	1.77	1.77	1.79	1.81	1.83
0.40	1.78	1.79	1.79	1.80	1.81	1.84	1.86
0.45	1.79	1.79	1.80	1.80	1.81	1.84	1.86
0.50	1.78	1.78	1.79	1.79	1.80	1.82	1.84
0.55	1.76	1.76	1.76	1.77	1.77	1.79	1.80
0.60	1.73	1.73	1.73	1.74	1.74	1.75	1.75
0.65	1.71	1.72	1.72	1.72	1.72	1.72	1.72
0.70	1.73	1.73	1.73	1.74	1.74	1.74	1.74
0.75	1.82	1.82	1.82	1.83	1.83	1.84	1.85
0.80	2.01	2.01	2.02	2.02	2.04	2.07	2.09
0.85	2.31	2.32	2.33	2.35	2.37	2.43	2.48
0.90	2.75	2.77	2.78	2.81	2.85	2.95	3.02
0.95	3.31	3.34	3.36	3.40	3.46	3.61	3.71
1.00	4.00	4.04	4.07	4.12	4.20	4.40	4.55

TABLE 1 (Continued)

$k \backslash n$	$\infty$	40	25	18	$p=4$ 14	11	9	8	7
0.00	1.88	1.88	1.88	1.89	1.90	1.91	1.94	1.96	1.99
0.05	1.87	1.87	1.87	1.88	1.89	1.90	1.93	1.95	1.98
0.10	1.84	1.84	1.85	1.85	1.86	1.87	1.89	1.91	1.94
0.15	1.80	1.81	1.81	1.81	1.82	1.83	1.85	1.86	1.88
0.20	1.77	1.77	1.77	1.77	1.78	1.79	1.80	1.81	1.83
0.25	1.74	1.74	1.74	1.75	1.75	1.76	1.77	1.78	1.79
0.30	1.73	1.74	1.74	1.74	1.75	1.76	1.77	1.78	1.80
0.35	1.76	1.76	1.76	1.77	1.78	1.79	1.81	1.82	1.85
0.40	1.81	1.82	1.82	1.83	1.84	1.86	1.88	1.91	1.94
0.45	1.89	1.90	1.90	1.91	1.93	1.95	1.98	2.02	2.07
0.50	1.98	1.98	1.99	2.00	2.02	2.05	2.09	2.13	2.19
0.55	2.06	2.07	2.07	2.09	2.11	2.14	2.18	2.22	2.29
0.60	2.12	2.12	2.13	2.14	2.16	2.19	2.24	2.27	2.34
0.65	2.14	2.15	2.15	2.16	2.18	2.20	2.23	2.27	2.32
0.70	2.13	2.13	2.14	2.14	2.15	2.16	2.18	2.20	2.22
0.75	2.10	2.10	2.10	2.10	2.11	2.11	2.11	2.11	2.11
0.800	2.12	2.12	2.12	2.12	2.12	2.12	2.12	2.13	2.13
0.825	2.19	2.19	2.19	2.20	2.20	2.21	2.22	2.24	2.27
0.850	2.32	2.32	2.33	2.33	2.35	2.37	2.41	2.45	2.53
0.875	2.52	2.53	2.54	2.56	2.59	2.64	2.71	2.79	2.93
0.900	2.82	2.83	2.85	2.88	2.93	3.01	3.14	3.26	3.48
0.925	3.21	3.23	3.26	3.31	3.38	3.50	3.68	3.86	4.16
0.950	3.71	3.73	3.77	3.84	3.93	4.10	4.35	4.58	4.98
0.975	4.30	4.34	4.39	4.48	4.60	4.82	5.13	5.43	5.94
1.000	5.00	5.04	5.11	5.22	5.38	5.64	6.04	6.40	7.03
$k \backslash n$	$\infty$	40	25	18	$p=5$ 14	11	9	8	7
0.00	1.88	1.88	1.88	1.89	1.90	1.91	1.94	1.96	1.99
0.05	1.89	1.89	1.90	1.91	1.92	1.94	1.96	1.98	2.03
0.10	1.93	1.94	1.94	1.95	1.97	1.99	2.02	2.05	2.11
0.15	1.99	1.99	2.00	2.01	2.03	2.05	2.10	2.14	2.21
0.20	2.04	2.04	2.05	2.06	2.08	2.11	2.16	2.21	2.29
0.25	2.07	2.08	2.09	2.10	2.11	2.15	2.19	2.24	2.32
0.30	2.08	2.08	2.09	2.10	2.12	2.14	2.18	2.22	2.29
0.35	2.06	2.06	2.07	2.08	2.09	2.11	2.14	2.16	2.20
0.40	2.02	2.03	2.03	2.04	2.04	2.06	2.08	2.09	2.12
0.45	2.00	2.00	2.00	2.01	2.02	2.03	2.05	2.07	2.10
0.50	2.00	2.01	2.01	2.02	2.03	2.06	2.09	2.13	2.21
0.55	2.06	2.07	2.08	2.09	2.12	2.16	2.23	2.30	2.46
0.60	2.18	2.19	2.20	2.23	2.26	2.33	2.43	2.55	2.78
0.65	2.32	2.33	2.35	2.38	2.43	2.51	2.64	2.78	3.05
0.70	2.45	2.46	2.48	2.51	2.56	2.64	2.77	2.90	3.16
0.75	2.51	2.52	2.54	2.56	2.59	2.65	2.74	2.82	2.99
0.800	2.50	2.51	2.51	2.52	2.53	2.54	2.55	2.56	2.56
0.825	2.49	2.49	2.49	2.49	2.49	2.48	2.46	2.45	2.41
0.850	2.50	2.50	2.50	2.50	2.49	2.49	2.48	2.48	2.53
0.875	2.59	2.59	2.60	2.60	2.62	2.64	2.72	2.82	3.12
0.900	2.82	2.83	2.85	2.88	2.94	3.05	3.28	3.57	4.24
0.925	3.25	3.27	3.32	3.39	3.52	3.76	4.21	4.72	5.84
0.950	3.91	3.95	4.04	4.17	4.38	4.79	5.50	6.27	7.89
0.975	4.82	4.90	5.02	5.23	5.54	6.14	7.14	8.21	10.38
1.000	6.00	6.11	6.28	6.57	7.00	7.81	9.15	10.54	13.35



TABLE 2

$k \backslash p$	1	2	$\rho p 0(k)$ 3	4	5
1.000	2.00	3.00	4.00	5.00	6.00
1.025	2.04	3.15	4.39	5.80	7.45
1.050	2.08	3.31	4.81	6.71	9.18
1.075	2.11	3.47	5.26	7.72	11.21
1.100	2.15	3.64	5.74	8.85	13.56
1.125	2.19	3.82	6.25	10.08	16.2
1.150	2.23	4.00	6.80	11.44	19.3
1.175	2.27	4.18	7.37	12.92	22.8
1.20	2.31	4.37	7.97	14.5	26.6
1.25	2.38	4.76	9.27	18.1	35.7
1.30	2.46	5.17	10.71	22.3	47.0
1.35	2.54	5.60	12.27	27.1	60.5
1.40	2.62	6.05	14.0	32.6	76.8
1.45	2.70	6.52	15.8	38.9	96.2
1.50	2.78	7.01	17.8	45.9	119
1.55	2.86	7.51	19.9	53.7	146
1.60	2.95	8.03	22.2	62.4	176
1.65	3.03	8.57	24.7	72.1	212
1.70	3.11	9.12	27.3	82.9	253
1.75	3.19	9.69	30.1	94.7	299
1.8	3.27	10.3	33.1	108	352
1.9	3.44	11.5	39.5	137	480
2.0	3.61	12.8	46.8	173	640
2.1	3.77	14.2	54.8	214	839
2.2	3.94	15.6	63.6	262	1083
2.3	4.11	17.1	73.4	317	1379
2.4	4.28	18.7	84.0	381	1740
2.5	4.44	20.3	95.6	454	2160
2.6	4.61	22.0	108.2	536	2660
2.7	4.78	23.8	121.8	629	3250
2.8	4.95	25.7	136.5	732	3940
2.9	5.12	27.6	152.3	849	4740
3.0	5.29	29.5	169.3	978	5670
$k > 3$	$1.73k$	$3.35k^2$	$6.61k^3$	$13.1k^4$	$26.1k^5$
$n \backslash p$	1	2	$\Phi_p(n)$ 3	4	5
$\infty$	1.00	1.00	1.00	1.00	1.00
100	1.00	1.00	1.00	1.00	1.00
50	1.00	1.00	1.00	1.01	1.01
40	1.00	1.00	1.00	1.01	1.02
30	1.00	1.00	1.01	1.02	1.03
25	1.00	1.00	1.01	1.02	1.05
20	1.00	1.01	1.02	1.04	1.07
18	1.00	1.01	1.02	1.05	1.09
16	1.00	1.01	1.03	1.06	1.12
15	1.00	1.01	1.03	1.07	1.14
14	1.00	1.01	1.04	1.08	1.16
13	1.00	1.02	1.04	1.10	1.19
12	1.00	1.02	1.05	1.11	1.23
11	1.00	1.02	1.06	1.14	1.28
10	1.01	1.03	1.08	1.17	1.35
9	1.01	1.03	1.09	1.22	1.47
8	1.01	1.04	1.12	1.30	1.66
7	1.01	1.05	1.17	1.42	2.03

IV. VARIATION OF PROBABLE ERROR WITH  $\varepsilon$ 

Examination of Table 1 shows that, for values of  $|k| < 0.9$ ,  $\rho_{p0}(k, n)$  differs from  $\rho_{p0}(k, \infty)$  by at most 2 per cent. for  $n \geq 18$  and by at most 1 per cent. for

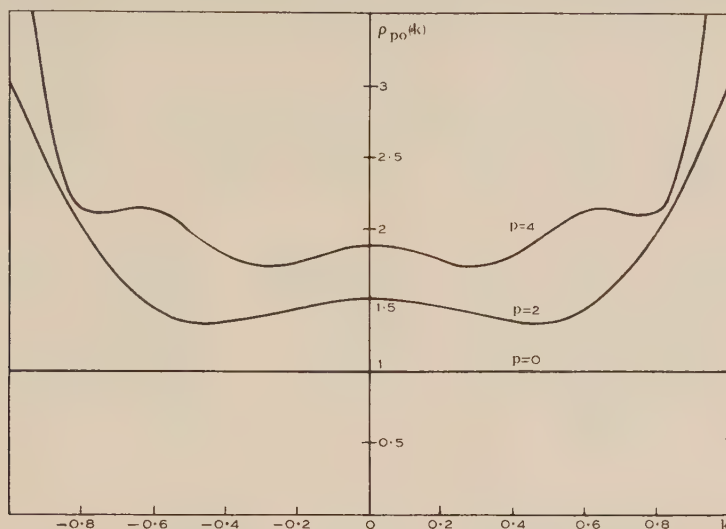


Fig. 1.—Variation of probable error in the region of interpolation ( $p$  even).

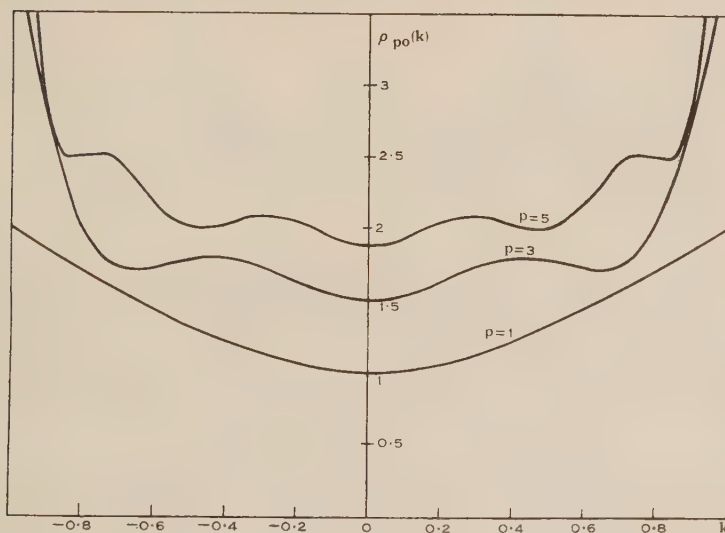


Fig. 2.—Variation of probable error in the region of interpolation ( $p$  odd).

$n \geq 25$ . Hence for reasonable values of  $n$  the curves giving the variation of the probable error in the region of interpolation are practically identical with the curves for  $n = \infty$ .

The curves for  $n = \infty$  are drawn in Figure 1 for  $p=0, 2, 4$ , and in Figure 2 for  $p=1, 3, 5$ . It will be observed that the probable error has  $p$  minima and

$p-1$  maxima, all symmetrically located about  $\varepsilon=0$ . The maxima and minima are shallow and the general trend is for the probable error to increase slowly as we proceed from the centre point; that is, successive maxima and minima are somewhat higher than the preceding ones. Beyond the last minimum the probable error increases quite rapidly; we have in effect a transition region between the region of interpolation and the region of extrapolation.

The curve for a given  $p$  lies between the curves for  $p-1$  and  $p+1$ , touching the former at points for which  $\tau_{p,0}$  vanishes and the latter at points for which  $\tau_{p+1,0}$  vanishes. However, since the variation of  $r_p$  with  $p$  depends on the particular problem, the relative positions of the curves have little significance.

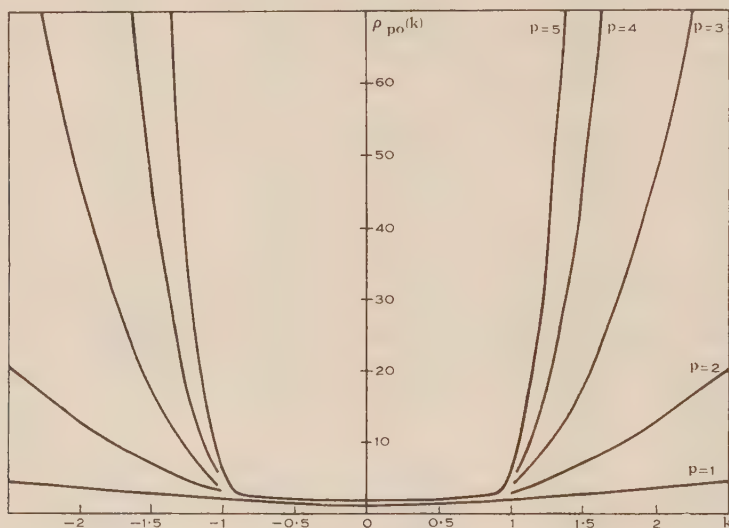


Fig. 3.—Variation of probable error in the region of extrapolation.

The curves of  $\rho_{p0}(k)$  in the region of extrapolation are sketched in Figure 3. It will be observed that the increase in the probable error is very rapid beyond  $k=1$ , especially for the higher values of  $p$ . The dangers attendant on the extrapolation of fitted polynomials have been stressed by Schultz and by Birge.

Table 2 extends to  $k=3$ . Beyond this value the function  $\rho_{p0}(k)$  can be written to a reasonable approximation as  $\alpha_p k^2$ , where  $\alpha_p$  is a constant.

#### V. LOCATION OF THE MAXIMA AND MINIMA

The maxima and minima of the weight function  $\rho_{p0}(k)$  are found to occur at the following values of  $k$ :

- (i)  $p=1$   $k=0$ .
- (ii)  $p=2$   $k=0$ , 0.447.
- (iii)  $p=3$   $k=0$ , 0.447, 0.655.
- (iv)  $p=4$   $k=0$ , 0.285, 0.655, 0.765.
- (v)  $p=5$   $k=0$ , 0.285, 0.469, 0.765, 0.830.

The maxima and minima of the probable error for a finite value of  $n$  will be displaced somewhat from these positions. The displacements are towards smaller values of  $k$  except for the last minimum in each degree. Since the values of  $k$  corresponding to the maxima and minima are complicated functions of  $v^2$ , it is simplest to tabulate these values for representative values of  $v^2$ . This is done in Table 3.

TABLE 3  
VALUES OF  $|k| (=2|\varepsilon|/n)$  FOR MAXIMA AND MINIMA

$v^2$	$n$	$p=5$				$p=4$			$p=3$		$p=2$
0.000	$\infty$	0.285	0.469	0.765	0.830	0.285	0.655	0.765	0.447	0.655	0.447
0.001	32	0.283	0.467	0.760	0.832	0.285	0.652	0.766	0.446	0.655	0.447
0.002	22	0.282	0.466	0.755	0.834	0.284	0.649	0.767	0.445	0.656	0.448
0.005	14	0.276	0.461	0.742	0.836	0.282	0.641	0.769	0.442	0.657	0.448
0.008	11	0.271	0.456	0.731	0.837	0.280	0.634	0.771	0.439	0.658	0.449
0.010	10	0.267	0.453	0.724	0.837	0.279	0.629	0.772	0.437	0.659	0.449
0.012	9	0.264	0.450	0.718	0.836	0.278	0.625	0.773	0.435	0.660	0.450
0.015	8	0.258	0.445	0.710	0.835	0.276	0.619	0.773	0.432	0.661	0.451
0.020	7	0.249	0.436	0.697	0.831	0.273	0.610	0.774	0.428	0.663	0.452

## VI. THE USE OF THE TABLES

The procedure for finding the probable error  $r[u_p(\varepsilon)]$  of the fitted value at the point  $\varepsilon$  may be summarized as follows :

(i) Evaluate  $|k| = 2|\varepsilon|/n$  to three decimal places,  $n$  being the number of observations.

(ii) (a) If  $|k| < 1$ , determine the value of  $\rho_{p0}(k, n')$  by interpolating between the listed values of  $k$  in Table 1 for the listed value  $n'$  closest to  $n$ . Estimate  $\rho_{p0}(k, n)$  by examining the variation with  $n$  at this point of the table.

(b) If  $|k| > 1$ , determine  $\rho_{p0}(k)$  by interpolation in Table 2, and  $\varphi_p(n)$  from the lower section of this table.

(iii) (a) If  $|k| < 1$ ,  $r[u_p(\varepsilon)] = \rho_{p0}(k, n) \cdot r_p/n^{\frac{1}{2}}$ ,

(b) If  $|k| > 1$ ,  $r[u_p(\varepsilon)] = \rho_{p0}(k) \cdot \varphi_n \cdot r_p/n^{\frac{1}{2}}$ ,

where  $r_p$  is the probable error of an observation of unit weight.

*Example 1.*—Consider the evaluation of the mechanical equivalent of heat from a series of observations at different temperatures (Birge 2). The particular values are here  $n=19$ ,  $\varepsilon=-5$ ,  $r_4=0.185$ .

Then  $|k|=10/19=0.526$ . From Table 1,  $\rho_{40}(0.50, 18)=2.00$ ,  $\rho_{40}(0.55, 18)=2.09$ . Thus  $\rho_{40}(0.526, 18)=2.00+0.26 \times 2 \times 0.09=2.05$ .

By inspection of the differences between the values at  $n=25$  and  $n=18$ ,  $\rho(19)$  may be taken to be equal to  $\rho(18)$ . Hence

$$r[u_4(-5)] = 2.05 \times 0.185/19^{\frac{1}{2}} = 0.087.$$

This is identical with the value obtained by Birge.

*Example 2.*—This is Birge's Illustrative Problem(2). Here  $n=7$ ,  $p=4$ , and  $r_4=7.25 \times 10^{-2}$ . The calculations are given in Table 4. The values calculated by Birge are shown in the last column for comparison.



TABLE 4

$\varepsilon$	$k$	$\rho_{40}(k, 7)$	$r[u_4(\varepsilon)] \times 10^2$	Birge's Value $\times 10^2$
0	0.000	1.99	5.45	5.46
1	0.286	1.80	4.93	4.91
2	0.571	2.31	6.33	6.36
3	0.857	2.64	7.23	7.20
		$\rho_{40}(k)$	$\varphi_4(7)$	
4	1.143	11.17	1.42	43.4
5	1.429	36.5	142	143
6	1.714	86.2	336	340

In many cases only a rough approximation to the probable error is required. Table 5 gives the probable error to 20 per cent. within the ranges of values of  $\varepsilon$  and  $n$  indicated.

TABLE 5

Degree $p$	Probable Error $r[u_p(\varepsilon)]$	Range of $ \varepsilon $	Range of $n$
1	$1.2r_p/n^{\frac{1}{2}}$	0 to 0.33n	$\infty$ to 7
2	$1.5r_p/n^{\frac{1}{2}}$	0 to 0.40n	$\infty$ to 7
3	$1.8r_p/n^{\frac{1}{2}}$	0 to 0.40n	$\infty$ to 7
4	$2.0r_p/n^{\frac{1}{2}}$	0 to 0.43n	$\infty$ to 7
5	$2.2r_p/n^{\frac{1}{2}}$	0 to 0.43n	$\infty$ to 10

## VII. CONCLUSION

It is apparent that the accuracy obtained for the probable error evaluated by use of the tables is more than adequate. The tables allow a rapid evaluation of the probable error to be made.

The tables are specially useful in the estimation of the probable error at points where  $V_j(\varepsilon)$  is not known; i.e. at points other than the points of observation. The method of Birge is to first calculate the  $V_j(\varepsilon)$  at these points. However, with tables of  $\rho_p(n, k)$  available the necessity for the calculation of the  $V_j(\varepsilon)$  disappears and the values of  $u_p(\varepsilon)$  can be obtained more quickly from the power-series representation.

## VIII. REFERENCES

- (1) SCHULTZ, H.—*J. Amer. Statist. Ass.* **25**: 139 (1930).
- (2) BIRGE, R. T.—*Rev. Mod. Phys.* **19**: 298 (1947).
- (3) FISHER, R. A., and YATES, F.—“Statistical Tables for Biological, Agricultural and Medical Research.” (Oliver and Boyd: Edinburgh, 1948.)
- (4) ANDERSON, R. L., and HOUSEMAN, E. E.—*Iowa St. Coll. Agric. Exp. Sta. Res. Bull. No. 297* (1942).
- (5) VAN DER REYDEN, D.—*Onderstepoort J. Vet. Sci.* **18**: 355 (1943).
- (6) ALLEN, F. E.—*Proc. Roy. Soc. Edinb.* **50**: 310 (1929-30).
- (7) BIRGE, R. T.—*Rev. Mod. Phys.* **1**: 1 (1929); **13**: 233 (1941); *Amer. J. Phys.* **13**: 63 (1945).

# REMARKS ON THE LATITUDE EFFECT OF COSMIC RAYS AT SEA-LEVEL

By H. D. RATHGEBER\*

[*Manuscript received December 23, 1949*]

## *Summary*

During a recent voyage from Australia to Japan the intensity of cosmic rays was measured with several Geiger counter telescopes of different designs. The conclusions to be drawn from these and other results are presented in this paper. It is shown that the latitude effect at sea-level (averaged over all azimuths) is approximately proportional to  $\cos^2 1 \cdot 2 \zeta$ , where  $\zeta$  is the zenith angle of the incident radiation, and that the vertical latitude effect is 18 per cent. In general, the latitude effect depends both on zenith angle and azimuth. The phenomenological relationships of these intensity variations with the zenith angle distribution at two fixed locations, i.e. at high latitudes and at the equator, and with the east-west effect at the equator, are deduced. After correction for longitude and zenith angle, Geiger counter telescope and ionization measurements do not differ by more than their errors of measurements; the average specific ionization does not vary with latitude. Qualitative explanations of the equality of the latitude effects for mesons and for electrons, as well as of the latitude effect of small extensive showers are given.

## I. INTRODUCTION

Recently, the intensity of cosmic rays was measured with several Geiger counter telescopes of different designs during a voyage from Australia to Japan and back. The geometrical arrangement, other relevant data, and the results are extensively described by Law, McKenzie, and Rathgeber(1). For the present it is sufficient to reproduce here a summary of their results.

An uncorrected latitude effect of  $20 \pm 1$  per cent. was measured with a narrow angle coincidence telescope both for the total vertical radiation and the vertical radiation filtered by 10 cm. of lead between the counters. For a wide angle coincidence telescope which subtended nearly the whole hemisphere the latitude effect reduces to  $13 \pm 1$  per cent. Further it has been shown that a latitude effect of some 10 per cent. exists for extensive showers of an average spread of 1 metre. An observed latitude effect of extensive showers of which one particle is able to penetrate 10 cm. of lead falls, however, within the standard deviation of the measurements.

Not all the results of Law, McKenzie, and Rathgeber(1) are in satisfactory agreement with previous measurements. Thus the latitude effect of 20 per cent. for the total vertical radiation exceeds by a factor two the generally accepted value of about 10 per cent.(2, 3). Since previous Geiger counter coincidence and ionization chamber measurements also differ among themselves by more than their errors of measurements an attempt has been made in this

\* Sir Thomas Lyle Research Fellow, Physics Department, University of Melbourne.

paper to analyse the factors influencing the latitude effect and to reduce the existing measurements to standard conditions.

In the recent experiment by Law, McKenzie, and Rathgeber(1) it was found that the latitude effects of the meson and electron components were very similar. This was interpreted as evidence that at sea-level the electrons are in equilibrium with the mesons. However, since the momentum spectrum of the primary cosmic rays varies with latitude, this interpretation can apply only to experiments at any one latitude and the observation receives further analysis below.

Further, the high probability of a latitude effect of small extensive showers, as observed by Law, McKenzie, and Rathgeber, calls for a re-examination of the process of their production. Different methods of energy transport by particles from the limit of the atmosphere to sea-level are considered and a qualitative answer outlined, which is not inconsistent with the cascade theory of showers.

## II. THE LATITUDE EFFECT OF THE TOTAL RADIATION

### (a) *The Variation of the Latitude Effect with Zenith Angle*

On the voyage reported(1), the latitude effect was measured simultaneously with several Geiger counter telescopes. One of these subtended an angle of  $26 \times 33^\circ$  in the vertical direction, another, also without absorber,  $90 \times 90^\circ$ . Latitude effects of 20 and 13 per cent., respectively, were obtained. Previous Geiger counter telescope and ionization chamber measurements show the same kind of difference. In addition, the Geiger counter telescope which subtends the whole hemisphere gave a result very near to that of spherical ionization chambers which are equally sensitive to radiation coming from all directions. These observations suggest that the latitude effect decreases for increasing zenith angles of the incoming radiation.

Variations in the temperature, and barometric effects for different zenith angles might have produced the differences in the latitude effect. This is excluded since earlier measurements with the same equipment(4) indicate that a difference of at most 1 per cent. could be introduced by these factors.

The different equipments used by different observers to measure the latitude effect integrate the effect over several ranges of zenith angle. Hence the relationship between latitude effect and zenith angle cannot be determined uniquely from the data, but it is possible to find a simple function which will satisfy the measurements.

If the cosmic ray intensity for a zenith angle  $\zeta$  is assumed(5) to be  $I_\zeta = I_0 \cos^2 \zeta$  which is a sufficient approximation for the present purpose, the following formulae for the flux coming from an annular region of width  $d\zeta$  at a zenith angle  $\zeta$  follow directly from Figure 1.

#### (i) *Spherical Ionization Chamber*

$$J_\zeta = I_{01} (\cos^2 \zeta) A_I \frac{1}{R^2} \int_{\alpha=0}^{2\pi} R \sin \zeta d\alpha \cdot R d\zeta, \\ = 2\pi I_0 A_I \sin \zeta \cos^2 \zeta d\zeta. \quad \dots\dots\dots (1)$$

$A_I$  is the effective area of cross section of the chamber.

(ii) *Circular Area with Vertical Normal*

$$J_{\zeta} = 2\pi I_0 A_C \sin \zeta \cos^3 \zeta d\zeta. \quad \dots\dots\dots (2)$$

As for the present case, the round sensitive area may be considered to be a sufficient approximation to a square one of the same size  $A_C$ , the formula obtained may be applied to the wide angle coincidence telescope (Law, McKenzie, and Rathgeber 1, count 3).

(iii) *Vertical Coincidence Telescope*

$$J_{\zeta} = 2\pi I_0 A_{\zeta} \sin \zeta \cos^2 \zeta d\zeta. \quad \dots\dots\dots (3)$$

$A_{\zeta}$  is the common area of the two outermost trays of Geiger counters for rays at an angle  $\zeta$ . It is shown crosshatched in Figure 1.  $A_{\zeta}$  has been evaluated

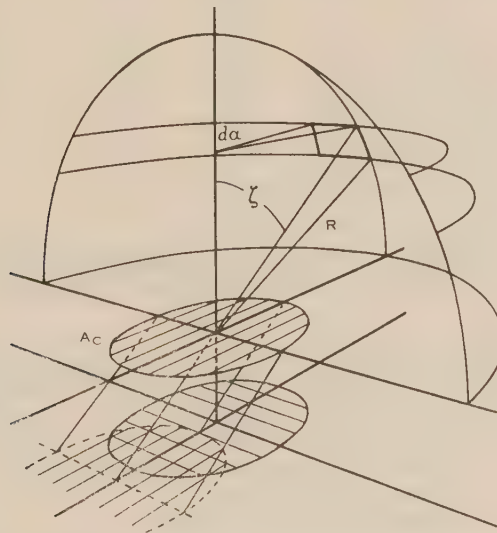


Fig. 1.—Diagram for the calculation of zenith angle sensitivities of cosmic ray instruments.

by numerical methods in terms of the zenith angle  $\zeta$  and the area  $A_C$  of one Geiger counter tray. As the same considerations as under (ii) apply here also, the formula may be used for the narrow angle coincidence telescopes (Law, McKenzie, and Rathgeber 1, counts 1 and 2).

(iv) The zenith angle sensitivity of the counter telescope used by Neher and Pickering(6) was also determined by numerical computations at a few points.

In Figure 2 the calculated zenith angle sensitivities have been plotted in such a way that the areas enclosed by the different curves are equal. At first sight it appears wrong that such an instrument as a spherical ionization chamber of which the sensitivity is the same in all directions should receive no radiation in the vertical direction and a maximum at a zenith angle of about  $35^\circ$ . This



can be understood by considering that the area from which radiation is received in the vertical direction tends to zero in the limit whilst it becomes a maximum for a zenith angle of  $90^\circ$ . The  $\cos^2$  law of cosmic radiation makes the zenith angle sensitivity decrease again as soon as the decrease in  $\cos^2$  exceeds the increase of the area of the annulus which is proportional to the sine.

A comparison of the zenith angle sensitivities and of the respective latitude effects shows that the latitude effect decreases only slightly for zenith angles up to  $30^\circ$  and then falls off rapidly for larger angles. Therefore some mathematically simple functions were assumed which do this, and the average latitude effects were calculated for the geometry of the different instruments. The results are summarized and compared with the measurements in Table 1. Whilst the  $\cos^2 1.2 \zeta$  function gives the best fit, the values for  $\cos^3 \zeta$  and  $\cos 1.5 \zeta$  also fall well within the limits of the errors.

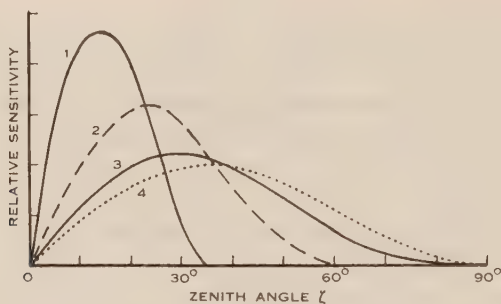


Fig. 2.—Relative sensitivities of:

1. Geiger counter telescope (Law, McKenzie, and Rathgeber 1, count 1).
2. Geiger counter telescope(6).
3. Geiger counter telescope (Law, McKenzie, and Rathgeber 1, count 3).
4. Spherical ionization chamber.

Several measurements of the latitude effect are compared in Table 2 before and after applying calculated corrections for the zenith angle variation.

#### *(b) Correction for Seasonal Effect of Cosmic Rays*

The measurements by Compton and Turner(7) which extended over several years have shown that the latitude effect increases by about 2 per cent. from summer to winter for ionization chamber measurements. The corrections which have to be applied for this effect are also shown in Table 2. The differences have been reduced, but it is quite evident that they still exceed the experimental error.

#### *(c) The Variation of the Latitude Effect with Longitude*

The existence of a variation of cosmic ray intensity at the geomagnetic equator with longitude was first observed by Clay(8) and subsequently confirmed by others(2, 3). In Figure 3 the values taken from Millikan and Neher's world

TABLE 1

CALCULATED RATIOS OF THE LATITUDE EFFECT AVERAGED FOR SEVERAL MEASURING INSTRUMENTS  
TO THE VERTICAL LATITUDE EFFECT

Latitude Effect in Function of Zenith Angle $L_{\zeta}=L_0 f(\zeta)$ $f(\zeta)$	Ratio of Averaged Latitude Effect to Vertical Latitude Effect				Ratio to Count 1		
	Count 1(1)	Neher and Pickering	Count 3(1)	Ionization Chamber	Neher and Pickering	Count 3(1)	Ionization Chamber
Cos .. ..	0.95	0.89	0.80	0.75	0.94	0.84	0.79
Cos 1.5 ..	0.89	0.77	0.56	0.51	0.86	0.63	0.57
Cos <sup>2</sup> ..	0.90	0.81	0.66	0.60	0.90	0.73	0.67
Cos <sup>2</sup> 1.2 ..	0.87	0.74	0.57	0.50	0.85	0.65	0.58
Cos <sup>2</sup> 1.5 ..	0.82	0.63	0.42	0.37	0.77	0.51	0.45
Cos <sup>3</sup> ..	0.86	0.73	0.57	0.50	0.85	0.66	0.58
Cos <sup>4</sup> ..	0.82	0.67	0.50	0.45	0.82	0.61	0.55
Measured ..					0.75	0.65	0.60
					$\pm 0.10$	$\pm 0.08$	$\pm 0.08$

TABLE 2

LATITUDE EFFECTS OF SEVERAL MEASURING INSTRUMENTS CORRECTED FOR APERTURE ANGLE AND  
SEASONAL EFFECT

Author	Measured Latitude Effect	Aperture Angle	Corrected Latitude Effect	Season of Measure- ment at High Latitude	Correction Seasonal Effect	Corrected Latitude Effect	Longitude at Equator
Johnson and Read(10) ..	$14 \pm 4$	$12 \times 22$	$15 \pm 4$	Summer	+1	$16 \pm 4$	85 W.
Neher and Pick- ering(6) ..	$15 \pm 1$	$20 \times 55$	$20 \pm 1$	No data			110 E.
	$10 \pm 1$	$20 \times 55$	$14 \pm 1$	No data			160 W.
Auger <i>et al.</i> (11)	$10 \pm 2$	$15 \times 48$	$13 \pm 2$	Summer	+1	$14 \pm 2$	85 W.
Morris, Swann, and Taylor(12)	$5 \pm 1$	$18 \times 18$	$6 \pm 1$	Summer	+1	$7 \pm 1$	35 W.
Law, McKenzie, and Rath- geber(1) ..	$20 \pm 1$	$26 \times 33$	$22 \pm 1$	Winter	-1	$20 \pm 1$	140 E.
	$13 \pm 1$	$90 \times 90$	$22 \pm 1$	Winter	-1	$20 \pm 1$	140 E.

survey with an ionization chamber(9) are plotted as curves, both as measured and divided with factors 0.50 and 0.60. The uncorrected and corrected data from Table 2 are also shown in the same graph. With the exception of one, all corrected measurements lie approximately on the curve obtained from the

ionization measurements with a factor 0.60. The remaining differences may well be attributed to the different meteorological conditions, and to the errors of reduction, each of which might amount to 2 per cent.

The difference between the observed factor 0.6 and the calculated one of 0.5 (Table 1) appears to exceed the errors of observation and calculation. It could be accounted for by a decrease of the specific ionization for inclined rays or by a slight decrease of the specific ionization from the knee to the equator.

In the above corrections the azimuthal variation of cosmic ray intensity, which for instance appears in the east-west effect, was not taken into consideration. Its influence in the measurements using square Geiger counter trays(1)

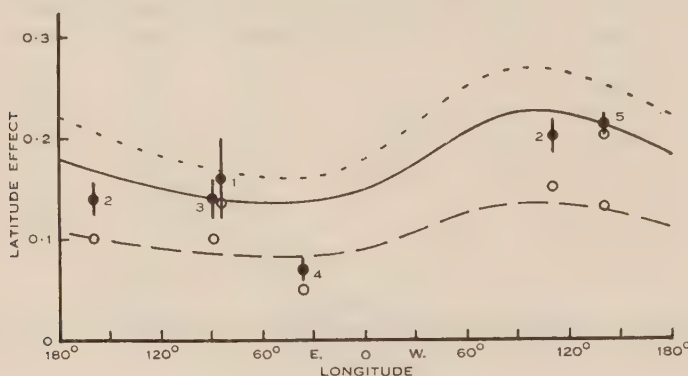


Fig. 3.—Longitude effect for vertical incidence of cosmic rays.

———— Ionization chamber (Millikan and Neher 9).

----- Ionization chamber divided by 0.50.

- · - · - Ionization chamber divided by 0.60.

○ Uncorrected Geiger counter measurements.

● Corrected for zenith angle and seasonal effects.

1. Johnson and Read(10). 2. Neher and Pickering(6). 3. Auger *et al.*(11). 4. Morris, Swann, and Taylor(12). 5. Law, McKenzie, and Rathgeber(1).

should not have exceeded the other errors. However, the coincidence telescope used by Neher and Pickering was asymmetrical in relation to azimuth. A distinct difference between latitude effects measured in the easterly and westerly directions at a zenith angle of  $45^\circ$  was observed by Johnson and Read(10). This could explain the difficulty of fitting Neher and Pickering's results to the calculated averages.

The otherwise successful reduction of coincidence telescope and ionization chamber measurements signifies that the ratio of the number of cosmic rays passing through a certain area to the ionization produced by them does not change appreciably with latitude. In other words the specific ionization of cosmic rays is independent of latitude within the errors of measurement.

As more than nine-tenths of the cosmic ray particles at sea-level possess energies above the energy of minimum specific ionization and their specific

ionization therefore increases only very slowly with energy, no conclusions about the respective energy spectra at high latitudes and at the equator can be drawn from this result.

Finally, it can be concluded from Figure 3 that the directly measured latitude effect for rays of vertical incidence averages  $18 \pm 2$  per cent. This value is not corrected for temperature effect, barometric effect, height of the meson producing layer(13), temperature of the stratosphere(14), or range of the primary cosmic ray particles(14).

The vertical latitude effect averaged in this way is equal, apart from the quoted corrections, to the undisturbed latitude effect(15, 16), i.e. the latitude effect which would be produced if the dipole equivalent to the earth's magnetic field would pass through the centre of the globe.

### III. THE ZENITH ANGLE VARIATION OF THE LATITUDE EFFECT AND THE ZENITH ANGLE DISTRIBUTION OF COSMIC RAY INTENSITY

It has been shown in the preceding section that the latitude effect averaged over all azimuths depends on the zenith angle at which it is measured. It must be concluded from this result that the zenith angle distribution of cosmic rays measured at a fixed location should vary with latitude, as will be shown.

If the latitude effect  $L_\zeta$  at a zenith angle  $\zeta$  is defined in the usual way(2, 3) the relationship

$$L_\zeta = (I_{\zeta K} - I_{\zeta M}) / I_{\zeta K} \dots\dots\dots (4)$$

is obtained with the cosmic ray intensities  $I_{\zeta K}$  and  $I_{\zeta M}$  measured at the knee and at the minimum of intensity near the geomagnetic equator respectively and at a zenith angle  $\zeta$ . Introducing  $G_{\zeta K}$  and  $G_{\zeta M}$  the zenith angle distributions of cosmic ray intensity by the definitions

$$I_{\zeta K} = I_{OK} G_{\zeta K},$$

and

$$I_{\zeta M} = I_{OM} G_{\zeta M},$$

the vertical intensity being  $I_O$ , and considering that  $I_{OM} = I_{OK}(1 - L_O)$  by definition, the following expressions are obtained :

$$L_\zeta = 1 - (1 - L_O) G_{\zeta M} / G_{\zeta K}, \dots\dots\dots (5)$$

$$G_{\zeta M} = G_{\zeta K} (1 - L_\zeta) / (1 - L_O). \dots\dots\dots (6)$$

A decrease of the latitude effect for increasing zenith angles requires that  $G_{\zeta M} / G_{\zeta K}$  increases, i.e. that the intensity decreases less rapidly at the equator than at the knee for increasing zenith angles. This has in fact been observed by Leprince-Ringuet and Auger(17). Further, their results show that the intensities for a zenith angle of  $60^\circ$  are equal at the equator and at Buenos Aires, i.e. the latitude effect becomes zero, as deduced before from Law, McKenzie, and Rathgeber's measurements.



The latitude effect has been measured by Johnson and Read(10) for inclined rays in the easterly and westerly directions. The latitude effects  $L_{\zeta_E}$  and  $L_{\zeta_W}$  are defined by

$$L_{\zeta_E} = \frac{I_{\zeta_K} - I_{\zeta_{ME}}}{I_{\zeta_K}}, \dots\dots\dots (7)$$

and

$$L_{\zeta_W} = \frac{I_{\zeta_K} - I_{\zeta_{MW}}}{I_{\zeta_K}}, \dots\dots\dots (8)$$

in terms of  $I_{\zeta_{ME}}$  and  $I_{\zeta_{MW}}$  which are the intensities in the easterly and westerly directions at the equator. As the east-west effect  $E_{\zeta_{EW}}$  is defined by

$$E_{\zeta_{EW}} = \frac{I_{\zeta_{MW}} - I_{\zeta_{ME}}}{\frac{1}{2}(I_{\zeta_{MW}} + I_{\zeta_{ME}})}, \dots\dots\dots (9)$$

the introduction of (7) and (8) leads to

$$E_{\zeta_{EW}} = \frac{L_{\zeta_E} - L_{\zeta_W}}{1 - \frac{1}{2}(L_{\zeta_E} + L_{\zeta_W})}.$$

Using the values  $L_{45E} = 16$  and  $L_{45W} = 4$  per cent. obtained in the measurements mentioned above(10) an east-west effect

$$E_{45EW} = \frac{12\%}{0.9} = 13 \pm 3\%$$

results for a zenith angle of  $45^\circ$ . Direct measurements give  $15 \pm 2\%$ (18).

The interrelations between the different effects shown here originate in the fact that these effects are only convenient relationships introduced by definition between the intensities at different localities and in different directions. Thus the interrelations do not have physical significance but are useful equations for checking some results against others.

#### IV. THE LATITUDE EFFECT OF MESONS AND ELECTRONS

In the measurements by Law, McKenzie, and Rathgeber(1) it was shown that the latitude effects of the total vertical radiation and of the vertical radiation filtered by 10 cm. of lead differed less than 1 per cent. This led to the conclusion that the latitude effect for the electron and meson components was the same within 2 per cent., thus confirming earlier measurements of lesser statistical accuracy(17).

More than 90 per cent. of the electrons at sea-level are created by the decay of mesons and by their collisions; less than 10 per cent. form part of showers from high energy primary electrons(19) which should not show any latitude effect. The influence of these falls within the errors of measurement and may be neglected.

The equality of the latitude effect of mesons and electrons has been explained by their being in equilibrium, but since the energy spectrum of the primaries varies with latitude this concept requires closer analysis. In calculations of the number of collisions and decay electrons Rossi and Klapman(20) have found

that to a first approximation their number per meson does not depend on the energy of the producing meson. Thus even if the energy spectrum of mesons near sea-level at the equator, which has not been measured, differs from the one at high latitudes the ratio of electrons to mesons should be the same at all latitudes. It is therefore not necessary to have recourse to another explanation assuming the equality of the meson spectrum near sea-level at all latitudes.

## V. THE LATITUDE EFFECT OF EXTENSIVE SHOWERS

Extensive showers observed at sea-level are considered to be produced by high energy primary electrons. A cascade process initiated by an electron or photon at the limit of the atmosphere can only reach sea-level if the energy of the primary particle is greater than  $10^{12}$  eV. Since charged particles of an energy greater than  $6 \times 10^{10}$  eV. can reach any point on the surface of the earth from any angle, the intensity of particles of energies higher than  $10^{12}$  is the same all over the earth. Hence the intensity of extensive showers should not vary with latitude.

The mean of the recalculated latitude effect of Neher and Pickering(21) and of Law, McKenzie, and Rathgeber(1) is  $9.5 \pm 3.5$  per cent. The probability of such a result occurring by chance is about 0.01. The barometric effect was eliminated from both measurements by multiple correlation(1). Any effect due to the different density of the atmosphere at the equator and in moderate latitudes may also be excluded by the observation that the shower intensity appears to depend on the barometric pressure but not on air temperature(22). It is concluded, therefore, that part of the primary radiation producing showers at sea-level is deflected by the earth's magnetic field and has a momentum smaller than 60 BeV/c.

It is interesting to consider the implications of this result. If the greatest part of the extensive shower rate is produced by mesons and their collision electrons, the rates for extensive and penetrating showers should be in a ratio of 2, the collision electrons being absorbed in one of the two trays. A ratio of  $(115-9)/(34-6) = 3.8$  has been observed by Law, McKenzie, and Rathgeber(1). This ratio, which agrees well with other results giving 1:4(23), excludes this explanation.

It may also be assumed that mesons produce in the atmosphere sufficient numbers of decay and collision electrons to produce in their turn extensive showers in the atmosphere. In the two measurements, the energy loss of electrons in the material of the deck and of the equipment is much smaller than the critical energy of 100 MeV. in air and may therefore be neglected. An electron of the critical energy would be subject to a Coulomb scattering of about 100 metres for one radiation length of 330 metres(24). Assuming that mesons decay into one electron and one neutrino, which is a sufficient approximation, we find that mesons with energies above 400 MeV. will be able to produce 1 metre showers. Considering that charged particles of vertical incidence from 3 to 15 BeV. are subject to the latitude effect it becomes evident that some of the 1 metre showers should be latitude dependent for this reason alone. Whether

the whole latitude effect of 1 metre showers can be accounted for by this hypothesis can only be shown in a detailed calculation.

The probability of meson decay is inversely proportional to their energy. The spectrum of the decay electrons will decrease therefore much more rapidly than the spectrum of the mesons which is proportional to  $p^{-\gamma}$  with  $\gamma$  between 2.5 and 3.5. From the analysis of big extensive showers by Molière(24) it is known that the spectrum of the electrons producing these showers is the same as that of the mesons. Therefore the big extensive showers are not produced by decay electrons. The same calculations show that more small extensive showers are observed than are expected from the primary electron spectrum. This excess of small extensive showers might also be explained by the same process as has been proposed here to account for the latitude effect of small extensive showers, namely their production by decay and collision electrons.

## VI. REFERENCES

- (1) LAW, P. G., MCKENZIE, C. D., and RATHGEBER, H. D.—*Aust. J. Sci. Res. A* **2** : 495 (1949).
- (2) JÁNOSY, L.—“Cosmic Rays.” (Oxford, 1948.)
- (3) MONTGOMERY, D. J. X.—“Cosmic Ray Physics.” (Princeton, 1949.)
- (4) CARO, D. E., LAW, P. G., and RATHGEBER, H. D.—*Aust. J. Sci. Res. A* **1** : 261 (1948).
- (5) GREISEN, K.—*Phys. Rev.* **61** : 212 (1942).
- (6) NEHER, H. V., and PICKERING, W. H.—*Phys. Rev.* **53** : 111 (1938).
- (7) COMPTON, A. H., and TURNER, R. N.—*Phys. Rev.* **52** : 799 (1937).
- (8) CLAY, J.—*Proc. Acad. Sci. Amst.* **33** : 711 (1930).
- (9) MILLIKAN, R. A., and NEHER, H. V.—*Phys. Rev.* **50** : 15 (1936).
- (10) JOHNSON, T. H., and READ, D. N.—*Phys. Rev.* **51** : 557 (1937).
- (11) AUGER, P., GRÉGOIRE, R., MAZE, R., and GOLDSCHMIDT, B.—*C.R. Acad. Sci. Paris* **209** : 794 (1939).
- (12) MORRIS, P. A., SWANN, W. F. G., and TAYLOR, H. C.—*Phys. Rev.* **74** : 1102 (1948).
- (13) RATHGEBER, H. D.—*Nature* **162** : 303 (1948).
- (14) DUPÉRIER, A.—*Nature* **163** : 369 (1949).
- (15) BRUINS, E. M.—*Physica* **2** : 879 (1935).
- (16) VALLARTA, M. S.—*Phys. Rev.* **74** : 1837 (1948).
- (17) LEPRINCE-RINGUET, L., and AUGER, P.—*J. Phys. Radium* **5** : 193 (1934).
- (18) JOHNSON, T. H.—*Rev. Mod. Phys.* **10** : 193 (1938).
- (19) ROSSI, B., and GREISEN, K.—*Phys. Rev.* **61** : 121 (1942).
- (20) ROSSI, B., and KLAPMAN, S. J.—*Phys. Rev.* **61** : 414 (1942).
- (21) NEHER, H. V., and PICKERING, W. H.—*Phys. Rev.* **58** : 665 (1940).
- (22) AUGER, P., and DAUDIN, J.—*Phys. Rev.* **61** : 91 (1942).
- (23) AUGER, P., MAZE, R., EHRENFEST, P. JR., and FRÉON, A.—*J. Phys. Radium* **10** : 39 (1939).
- (24) MOLIERE, G.—“Cosmic Radiation.” p. 26. (New York, 1948.)



# THE FORMATION OF RAIN BY COALESCENCE

By E. G. BOWEN\*

[*Manuscript received March 2, 1950*]

## *Summary*

It is generally acknowledged that drizzle or light rain can fall from clouds which do not reach freezing level and cases have recently been described in which moderate to heavy rain has been observed to fall from such clouds. A simple theory is developed to account for the phenomenon, based on the initial growth of cloud droplets by condensation followed by the growth of a small fraction of their number by coalescence. These grow in their ascent through the cloud until they are large enough to remain in suspension in the upward air current, after which they fall as rain. It is shown that for a given set of cloud conditions the maximum height reached by the drops increases with increasing vertical air velocity and that the size of the drops emerging from the base of the cloud is nearly a linear function of the height attained. The time for the precipitation to appear, on the other hand, is an inverse function of the upward air velocity.

Experimental observations of rain from non-freezing clouds have distinguished two main types. The first of these shows an increase in drop diameter or rainfall intensity downward through the cloud, as would be expected if the drops followed a variety of trajectories within the cloud. The second type is one in which the drop trajectories tend to coincide, in which case there would be a maximum in the raindrop density and the rain water content at some defined height within the cloud. This has been verified qualitatively by radar observations and flight experiments.

## I. INTRODUCTION

In 1933 Bergeron(1) postulated a mechanism of rain formation in which raindrops began their life as ice crystals and then grew rapidly at the expense of surrounding water droplets. It is now generally accepted that this process plays an important part in the formation of rain from clouds which extend well above the freezing level. At the same time it is recognized that other mechanisms are possible and that drizzle or light rain can fall from clouds which consist wholly of water droplets. In considering this phenomenon Findeisen(2) showed that drops of the required size were not likely to form by condensation alone but could form if coalescence of cloud droplets was taking place. He calculated the rate of growth by coalescence and showed that drizzle would be expected to form in clouds of moderate thickness and large raindrops in clouds of greater thickness. As evidence for the fall of moderate or heavy rain from non-freezing clouds did not exist at that time, he came to the conclusion that coalescence occurred with small drops but not with larger ones. Houghton(3) made calculations of a similar nature but did not give any evidence for or against the process.

Since that time several accounts(4-6) have appeared of moderate to heavy rain having been seen to fall from clouds which did not reach freezing level,

\* Division of Radiophysics, C.S.I.R.O., University Grounds, Sydney.



suggesting that the mechanism by which rain forms in these clouds is worthy of further investigation. In the present paper an attempt is made to explain the production of rain from non-freezing clouds on the assumption that after the initial growth of cloud droplets by condensation, a small proportion of their number grows further by coalescence. Calculations of the rate of growth are made, similar to those of Findeisen and Houghton, and these are used to determine the paths taken by individual drops within a cloud. The effect of varying the cloud parameters is investigated and it is found that the final size of the raindrops formed by coalescence depends mainly on the vertical air velocity in the cloud. Their size is less affected by variations in the cloud water content and the degree of supersaturation. It is also deduced that under appropriate conditions a concentration of drops may occur within a relatively narrow range of heights above the cloud base. This effect has been observed in both flight and radar experiments.

## II. THEORY

Consider the course of events in a single column in the atmosphere in which the air is ascending with uniform vertical velocity and with no horizontal motion. It is known that the air cools adiabatically as it ascends and that cloud droplets form when it becomes saturated. The droplets then grow by condensation at a rate which is determined mainly by the degree of supersaturation of the air. If the mean upward air velocity is  $v$ , the motion of any one cloud droplet is given by

$$\frac{dh}{dt} = v - u, \quad \dots\dots\dots (1)$$

where  $h$  = height above the base of the cloud,

$t$  = time,

$u$  = terminal velocity of the cloud droplet.

Stokes's law will hold for cloud droplets of the size we are considering and their terminal velocity is given by

$$u = \frac{g\rho D^2}{18n}, \quad \dots\dots\dots (2)$$

where  $g$  = acceleration due to gravity,

$D$  = drop diameter,

$n$  = viscosity of air,

$\rho$  = density of water.

The full expression for growth by condensation has been given by Frössling(7) but it is sufficient for the present approximate discussion to use a much simpler one derived from that given by Houghton(8):

$$D^2 = 8kQ \cdot \Delta q \cdot t, \quad \dots\dots\dots (3)$$

where  $k$  = diffusion coefficient of water vapour in air,

$Q$  = saturated vapour density,

$\Delta q$  = percentage supersaturation.

Inserting (2) and (3) in equation (1), gives for the motion of the cloud droplet

$$\frac{dh}{dt} = v - \frac{8kg}{18n} \rho \cdot Q \Delta q \cdot t. \quad \dots\dots\dots (4)$$

Integrating and inserting the condition  $h=0$  where  $t=0$ , we have

$$h = vt - \frac{2kg}{9n} \rho \cdot Q \Delta q \cdot t^2, \quad \dots\dots\dots (5)$$

which gives the height above the cloud base reached by a cloud droplet after the lapse of time  $t$ . For typical values of cloud parameters the expression  $\frac{2kg}{9n} \rho \cdot Q \Delta q \cdot t^2$  is small, indicating that the cloud droplets will lag very little behind the ascending air and continue to rise with it as long as the up current persists.

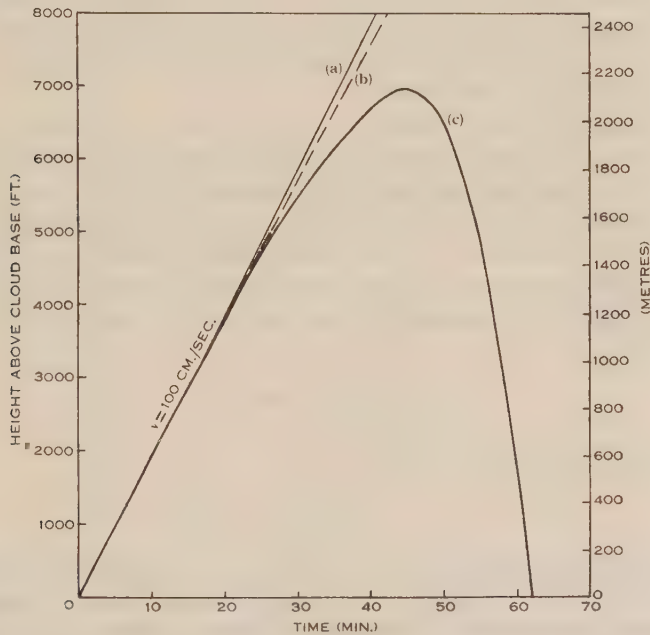


Fig. 1.—The motion of (a) the air, (b) cloud droplets, and (c) droplets which have grown by coalescence, in a cloud in which the mean upward air velocity is 100 cm./sec., average droplet diameter  $20\mu$  and cloud water content 1 g./m.<sup>3</sup>.

Taking as an example a cloud with a mean upward air velocity  $v=100$  cm./sec., a degree of supersaturation 0.1 per cent. and the usual value of the other constants at  $10^{\circ}\text{C}$ ., the motion of the cloud droplets with time is found to be as in curve (b) of Figure 1, the motion of the air being given by (a). It is evident that unless some mechanism of growth comes into action in addition to condensation, none of the droplets will grow large enough to fall against the upward air current in a cloud of reasonable thickness. If, however, as a consequence of the growth by condensation or due to turbulence and mixing, cloud

droplets of different diameter appear together in any small sample of cloud, they will fall relative to one another in the gravitational field and collisions will occur. If, on collision, coalescence takes place, rapid growth will follow.

The question whether water droplets coalesce on collision has been debated for many years. Rayleigh(9) is often quoted as showing that coalescence does not take place. His conclusions were drawn from experiments on the behaviour of drops in two impinging jets of tap water, so that they do not necessarily apply to droplets in clouds. He found that his drops did not normally coalesce in dust-free air, but he was at pains to point out that coalescence did take place remarkably easily in the presence of small quantities of contamination in the water, in the presence of many forms of dust, including atmospheric dust, or if there were small differences of potential between the drops. It is clear from his results that surface effects are crucial in determining whether coalescence takes place. In the absence of direct evidence as to whether cloud droplets coalesce on contact or not, it is proposed to consider the consequences which would arise if every collision in a cloud resulted in coalescence, and see whether the results are in agreement with observations.

Langmuir(10) has made a study of the conditions which obtain when a drop of diameter  $D$  falls through a cloud of slightly smaller droplets. On the above assumption that each collision results in coalescence, he derived a quantity  $E$ , the collection efficiency of the drop, or the fraction of droplets in its path which is picked up by coalescence. If the velocity of fall of the larger drop is  $u$ , the cloud has a liquid water content  $w$  grams per unit volume, and the rate of fall of the cloud droplets is negligible, then the rate of change of volume of the larger drop is

$$\frac{d. \text{ vol.}}{dt} = \frac{\pi D^2}{4} \cdot wuE, \dots\dots\dots (6)$$

and its rate of change of diameter is

$$\frac{dD}{dt} = \frac{wuE}{2}. \dots\dots\dots (7)$$

Langmuir showed that  $E$  increases from zero, or a very small value for a small drop, to nearly unity for a drop of raindrop diameter falling through a cloud of droplets. Some experimental checks of collection efficiency have been made in this Laboratory giving values in fair agreement with those of Langmuir: the latter have therefore been used in the calculations which follow.

Let us now trace the growth by coalescence of a drop in falling through a cloud which is assumed to have a mean upward air velocity  $v$  of 100 cm./sec., a cloud water content  $w$  of 1 g./m.<sup>3</sup> and an average droplet diameter of 20 $\mu$  after the initial growth period. If, due to turbulence or mixing, two droplets of slightly different diameter come together and coalesce, they will form a drop of approximately twice the mass. This will fall relative to the others. Its subsequent growth has been calculated from equation (7) above, using the values of  $E$  given by Langmuir, the terminal velocity  $u$  derived from Stokes's law for small droplets and that given by Laws(11) for larger drops. As both the terminal velocity and the collection efficiency change in a complex fashion with drop

diameter, the calculation has been made numerically. It is found that growth by coalescence is slow at first and only of the same order of magnitude as the growth by condensation. However,  $u$  and  $E$  increase rapidly with drop size and some time after collisions commence the drop is large enough to remain in suspension in the up current. On further growth it falls through the rising air, and, growing still further in its downward passage, it finally emerges as a raindrop from the base of the cloud.

The trajectory obtained in this way is given in Figure 1 (*c*), indicating that under the conditions specified the drop comes into equilibrium in the air current after 45 minutes at a height of 6950 ft. above the cloud base. It descends to the

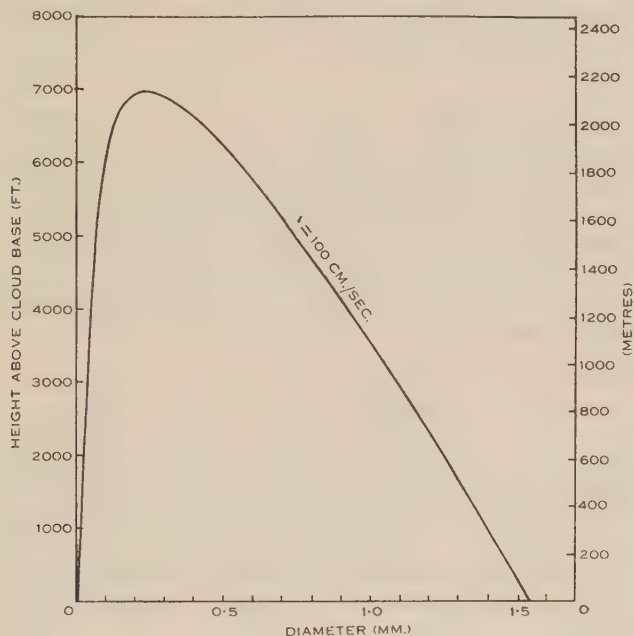


Fig. 2.—The variation of diameter with height of a drop growing by coalescence in a cloud in which the mean upward air velocity is 100 cm./sec., average droplet diameter  $20\mu$  and the cloud water content 1 g./m.<sup>3</sup>.

base in another 17 minutes, the whole process taking 62 minutes. The corresponding curve of height against diameter (Fig. 2) can then be obtained by a simple step, indicating that the final size of the drop is 1.55 mm.—a raindrop of moderate size.

### III. THE EFFECT OF VERTICAL AIR VELOCITY

Some early calculations showed that of the cloud parameters likely to affect the final drop size, vertical air velocity had a predominant effect. The effect of varying the vertical air velocity over a wide range of values has therefore been investigated and a series of curves similar to those of Figures 1 and 2 computed for mean rates of ascent  $v=10, 25, 50, 100$ , and 200 cm./sec. If condensation



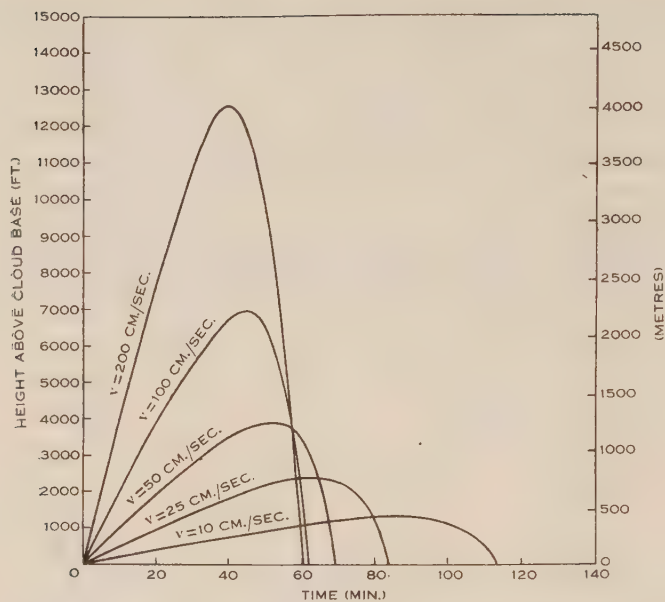


Fig. 3.—The trajectories of drops which grow by coalescence in clouds having a range of vertical air velocities from 10 to 200 cm./sec.

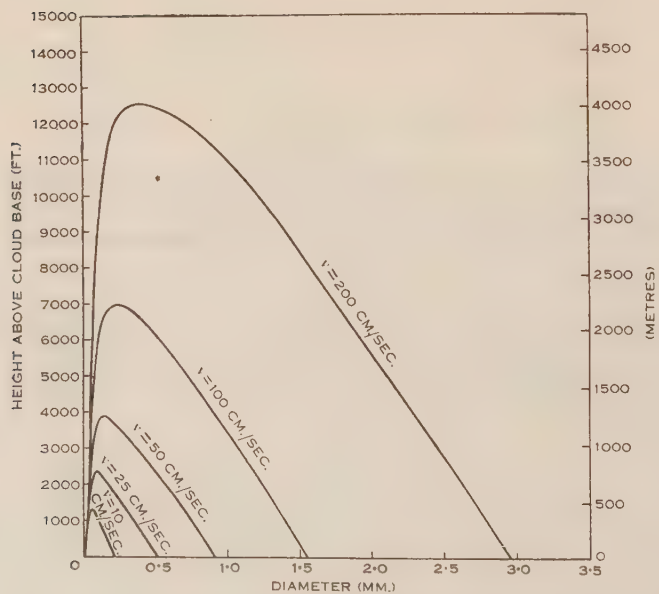


Fig. 4.—The change in drop diameter with height for a range of vertical air velocities.

took place on the same number of nuclei in each case, the degree of supersaturation attained through the bulk of the cloud would be approximately a linear function of the rate of ascent. It does not follow that this is true in natural clouds but it appears from Howell's(12) treatment of growth by condensation that it is in the right sense. As a first approximation, therefore, it is assumed that the percentage supersaturations corresponding to the above vertical air velocities are 0.01, 0.25, 0.5, 0.1, and 0.2 respectively. A series of height-time curves is then obtained as in Figure 3 and a corresponding series of curves of height against diameter as in Figure 4. It is seen that a whole range of raindrop sizes from 0.2 to 3 mm. is possible for upward air currents varying from 10 to 200 cm./sec. The general conclusion can therefore be drawn that a condensation-coalescence process of the type described is capable of accounting for raindrops of a wide range of sizes and could be the mechanism at work when rain is observed from non-freezing clouds.

It can also be seen from Figures 3 and 4 that if rain forms in the manner postulated, the maximum height reached by the drops will increase with the mean upward air velocity while the size of the drops emerging from the base of the cloud will, in turn, increase with the maximum height attained. The time for the whole process, on the other hand, will be an inverse function of the vertical air velocity, clouds with a low value of up current taking a long time to precipitate.

#### IV. THE EFFECT OF VARYING OTHER CLOUD PARAMETERS

The other cloud parameters which are likely to affect the trajectories of raindrops produced by this process are the degree of supersaturation, the cloud water content, and the average size of the cloud droplets. The effect of varying these quantities and of departures from Langmuir's values of collection efficiency has been investigated by repeating the calculations over a wide range of values. In each case it is found that, while the form of the trajectory might vary with changes in the parameters, the final drop size is relatively unchanged.

##### *(a) Degree of Supersaturation*

It follows from the treatment in Section II that the degree of supersaturation will have little direct effect on the final size of the drops since condensation contributes so little to their final mass. However, since the degree of supersaturation determines the initial rate of growth, it will influence the point at which growth by coalescence starts and in this way might affect the final drop size. Taking as an example the cloud conditions specified in Section II, it is found that doubling the degree of supersaturation causes the height of the trajectory to be reduced from 6950 ft. to 5100 ft. and the drop size from 1.55 mm. to 1.25 mm., that is, a change of 2:1 in percentage supersaturation causes a change of only 20 per cent. in the final drop size. It will be noted that the change is in the inverse sense, an increase in the degree of supersaturation causing a decrease in the height attained and a corresponding decrease in the size of the drop.

*(b) Cloud Water Content*

In the same way, the effect of changing the cloud water content has been calculated for a range of values from 0.5 gram per cubic metre to 1.5 grams per cubic metre, the other conditions being unchanged, namely, vertical air velocity  $v=100$  cm./sec., a degree of supersaturation of 0.1 per cent. and a mean cloud droplet size of  $20\mu$ . The curves of drop size against height shown in Figure 5

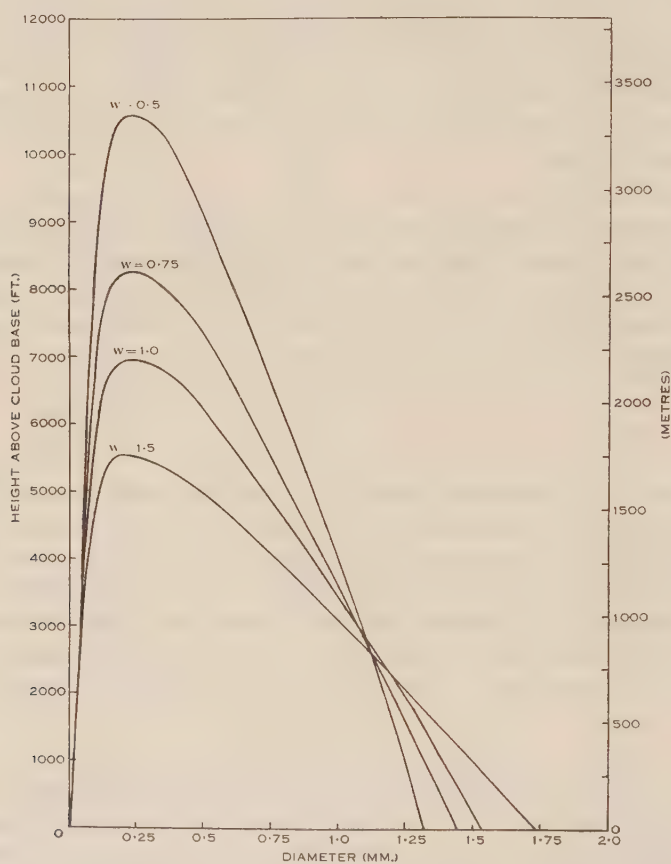


Fig. 5.—The change in drop diameter with height for different values of cloud water content.

are obtained, indicating that an increase in the ratio 3:1 in the cloud water content causes an increase of only 33 per cent. in the final drop size. This is due to the fact that, while the rate of growth increases with increased cloud water content, the length of the path over which growth takes place is reduced. It can be seen therefore that the final size of drops growing by coalescence is relatively insensitive to changes in cloud water content.

*(c) Collection Efficiency*

In the same way, the effect of departures from Langmuir's figures for collection efficiency has been computed for cloud conditions which are otherwise unchanged. This has been done by assuming that no growth by coalescence occurs until the cloud droplets attain a mean diameter of  $20\mu$ , after which growth takes place at fixed values of  $E$  ranging from 0.25 to 1. The curves of drop size against height obtained in this way are given in Figure 6, showing again that

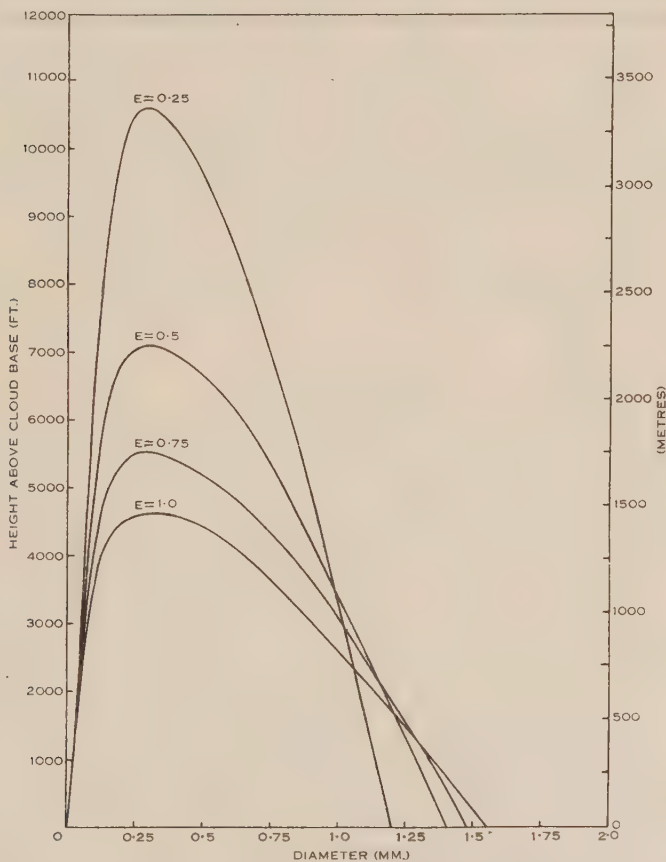


Fig. 6.—The change in drop diameter with height for different fixed values of "collection efficiency".

while the maximum height of the trajectory changes, the final size of the raindrops increases by only 30 per cent. for an increase of 4 : 1 in the value of  $E$ .

*(d) Cloud Droplet Size*

Cloud droplet size does not appear in the formula for growth by coalescence, but for a fixed cloud water content the effect of variation in the size of the cloud droplets appears as a change in the collection efficiency  $E$ . Calculations have been made for droplet diameters of 20, 30, and  $40\mu$ , giving the curves of Figure 7,



which again show a relatively small change in the final drop size. Calculations for cloud droplets smaller than  $20\mu$  are probably no longer valid due to uncertainties in the value of  $E$  but it is evident that the collision rate will fall off rapidly, which means that both the rate of growth and the fraction of cloud droplets growing by coalescence are much reduced when the cloud droplets are small. This might well be a critical condition for rain to form by the coalescence process. Knowledge of the collision process is not exact enough for firm conclusions to be drawn, however, but it is probable that the average droplet diameter in a cloud must grow to at least  $20\mu$  before coalescence becomes important.

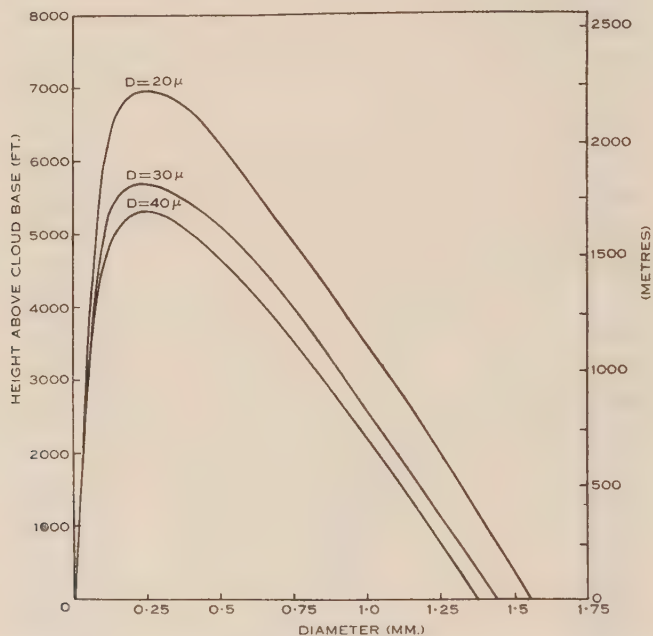


Fig. 7.—The change in drop diameter with height for clouds with an average cloud droplet diameter of 20, 30, and  $40\mu$  respectively.

## V. EXPERIMENTAL OBSERVATIONS

A complete experimental check of the theory would require measurement of the diameter and number of raindrops, the vertical air velocity and the cloud water content at different levels in typical clouds. Many of these measurements are difficult to carry out and it may be some time before they are accomplished. In the meantime, evidence in support of the theory has been obtained from observations of rain from non-freezing clouds by ground radar equipment and aircraft flights in the vicinity of Sydney, Australia. In general, it is found that such rain takes on two distinct forms, one in which the rain intensity increases gradually downwards through the cloud, the other in which there is a layer type of formation with a concentration of relatively large drops at some height above the cloud base.

The first of these may be regarded as the general case in which droplets start growing by coalescence at various levels in the cloud giving rise to a variety of trajectories of the type shown in Figure 3. It is known that the intensity of radio waves scattered from raindrops within a cloud is proportional to  $\Sigma ND^6$ , where  $N$  is the raindrop density and  $D$  the diameter of the drops. When a cloud of the above type is observed by radar, therefore, the raindrops would first be detected at that point in their descent where the quantity  $\Sigma ND^6$  is sufficient to give a measurable signal. Any subsequent increase in diameter of the raindrops as they fell through the cloud would give a further increase in signal intensity.

The second form, that in which a concentration of water drops is found at some height above the cloud base, shows a horizontal band structure when observed by radar. It is thought to be a particular case in which there is some factor at work tending to make all the raindrop trajectories coincide or nearly coincide. It will be shown that in these circumstances there would be a concentration of drops at the top of the trajectory.

A detailed description will now be given of the two types of non-freezing rain which have been observed.

#### *(a) The General Case*

A typical example of the first kind occurred on the morning of December 1, 1949, when light rain fell inland of the Radiophysics Laboratory. Maritime air was moving across the coast from an easterly direction with 6/8 cumulus cloud, the cloud base being at 2500 ft. An aircraft was operating overhead at the time and it was found that the cloud tops were generally at 10,000 ft. where the temperature was  $+2\frac{1}{2}^{\circ}\text{C}$ . A few heads appeared some 500 ft. higher but did not attain freezing level at any time.

Rain fell west of the Laboratory for several hours from 0930 hours onwards. It was observed on an SCR 717 radar operating on a wavelength of 10 cm. and modified to scan from horizon to horizon through the zenith. The sensitivity of the set was such that approximately  $10^4$  raindrops per cubic metre, each 0.5 mm. diameter, would just be detected if they filled the beam at a distance of 10,000 ft. The echo pattern observed at 1002 hours is shown in Plate 1 which is, in effect, a side elevation view in an east-west plane through the rain area. The observing point is at the centre of the baseline and the bright semi-circles are range markers at 5000, 10,000, and 15,000 ft. respectively. The rain echoes appear wholly on the right of the picture, the echoes on the left being from objects on the ground east of the observing point. The echo pattern changed only slowly during three hours' observation and was of the same general appearance throughout. In Plate 1 the rain echoes extend downwards from an average height of 6000 to 7000 ft. with the exception of the most distant column which extends from 10,000 ft. The gradation of echo intensity is not well reproduced in the plate, but a measure of echo intensity against height has been obtained by means of a microphotometer scanning vertically along the rain columns at the points marked (a), (b), and (c). The results are plotted in

Figure 8 in arbitrary units of echo intensity against height in feet. It is seen that there is satisfactory correspondence between them, all three showing an increase in echo intensity and therefore of rain intensity downward through the cloud. The rapid decrease of intensity below 2000 ft. is not significant as it is probably due to the shielding of the rain echoes near the ground by objects in the vicinity of the radar set.

The column structure shown in Plate 1 is very characteristic of this type of rain. The separate columns are probably related to individual convective cells in the manner described by Byers and Braham(13) for more active cumulus clouds.

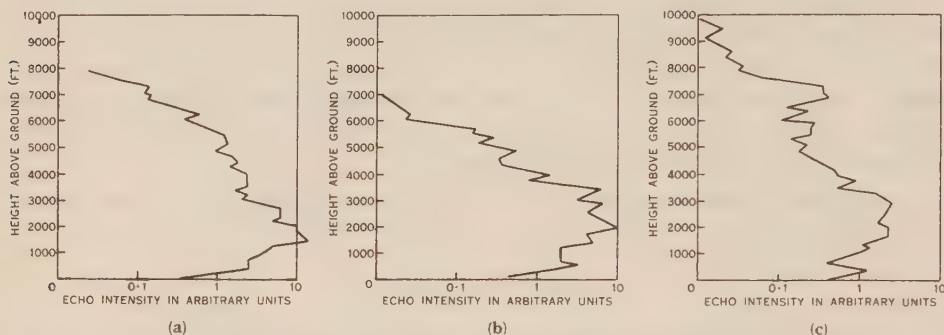


Fig. 8.—The radar echo intensity in arbitrary units plotted against height for three vertical sections (a), (b), and (c) through the rain area shown in Plate 1.

### (b) *The Horizontal Band Structure*

Considering now the particular case in which a concentration of drops appears at some level in the cloud, it is convenient first to estimate how the raindrop density and the rain water content would vary with height in a cloud in which the drop trajectories tend to coincide. Considering only those drops which are moving downward relative to the ground, if  $n$  is the number of raindrops of a given size crossing unit area per second, then  $N$ , the number of raindrops per unit volume at any level, is given by

$$N = \frac{n}{v - u} \dots \dots \dots (8)$$

This clearly increases upward from the base of the cloud and tends to infinity at the top of the trajectory. Numerical values of drop density have been calculated for the cloud conditions assumed in Section II, namely an upward air velocity of 100 cm./sec. and a cloud water content of 1 gram per cubic metre, giving the curve shown dotted in Figure 9. In the same way, the rain water content at any level is given by  $\frac{\pi N D^3}{6}$ , and this has also been computed in

terms of  $n$ , giving the values shown in the full line in Figure 9. Unlike raindrop density, the rain water content first falls off with increasing height, and then tends to infinity at the top of the trajectory due to the very great increase in raindrop density  $N$ .

Because of the assumption that all the drops have identical trajectories, the drops appear to be concentrated in a region of infinitesimal width and the rain water content appears to be infinite. In practice the trajectories are unlikely to coincide exactly and the region would have finite width and finite water content.

It would be expected, therefore, that, if a cloud were producing rain by the condensation-coalescence process and the raindrop trajectories tended to coincide, a concentration of water drops or an increase in rain water content

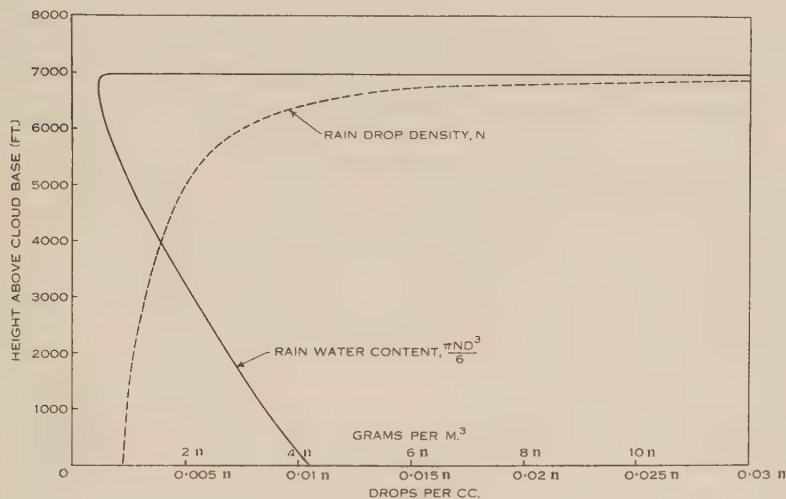


Fig. 9.—The variation with height of raindrop density and rain water content, both in terms of  $n$  the number of raindrops crossing unit area in unit time. The cloud conditions are as specified in Figure 1.

would occur at some height above the cloud base. This, in fact, has been observed by Smith(14) and the author in flights through cumulus clouds which were producing or about to produce rain. Some examples of this effect follow.

(i) *Flight Observations of June 1, 1948.*—On one of the occasions described by Smith, namely that of June 1, 1948, he was able to make measurements of raindrop size, drop density, and rain water content at a variety of levels. The instrument used for this purpose was an impactor in which a sensitive paper strip was moved at a uniform speed past a slot exposed to the air stream. The raindrops were recorded as spots on the paper and an estimate of the drop size obtained from the spot size by calibration. The minimum drop size capable of giving a record was approximately 0.1 mm. The raindrop diameters obtained are not claimed to be accurate, but the instrument is thought to give good comparative figures of drop density and water content at the different levels.

The meteorological conditions during the flight have been described in some detail by Smith. Light to moderate rain was falling from a cloud which approached, but did not reach, freezing level. The greatest height reached



by the cloud was 8400 ft. where the temperature was  $+\frac{1}{2}^{\circ}\text{C}$ . and the base of the cloud was at 1500 ft. Flights were made through it at 8000, 7000, and 6000 ft. and underneath the base at 1300 ft. No ice or snow particles were observed in the cloud and it was found that the rain water content at the top was considerably higher than elsewhere. Measurements made with the impactor gave the results shown in Table 1. They show that the mean drop diameter at the top of the cloud was about 0.3 mm. and was 0.7 mm. just below the cloud base, indicating that the drops grew as they fell through the cloud.

TABLE 1  
MEASUREMENTS OF RAINDROP DENSITY AND RAIN WATER CONTENT—JUNE 1, 1948

Height (ft.)	Estimated Diameter (mm.)	<0.2	0.2 to 0.4	0.4 to 0.6	0.6 to 0.8	0.8 to 1.0	1.0 to 1.2	1.2 to 1.4	1.4 to 1.6	Totals
8000	Drop density N/m. <sup>3</sup> .. ..	54,000	9,200	1,300	220	18	—	—	—	64,840
	Water content mg./m. <sup>3</sup> ..	25	130	85	40	7	—	—	—	287
7000	Drop density N/m. <sup>3</sup> .. ..	14,000	1,100	130	18	—	—	—	—	15,250
	Water content mg./m. <sup>3</sup> ..	7	15	8	3	—	—	—	—	33
6000	Drop density N/m. <sup>3</sup> .. ..	9,000	400	50	10	—	—	—	—	9,460
	Water content mg./m. <sup>3</sup> ..	5	6	3	2	—	—	—	—	16
1300	Drop density N/m. <sup>3</sup> .. ..	15,200	1,160	480	220	70	14	5	2	17,150
	Water content mg./m. <sup>3</sup> ..	7	16	31	40	27	10	6	4	141

The total water content at the different levels is plotted in Figure 10 and if a smooth curve is drawn through them it corresponds closely in form to that given by the theory in Figure 9. The total drop densities have been plotted in the same way, giving the dotted curve of Figure 10. This also agrees with that derived from the theory in showing a maximum in the upper part of the cloud, but it increases again towards the cloud base. It is thought that this is due to the fact that the drops grew in falling through the cloud and a certain number which were too small to be detected by the instrument at the top of the cloud had grown sufficiently for them to be counted at the base. Qualitatively, therefore, the observations agree with the theory in showing an increase in drop diameter as the drops fall through the cloud, and in showing both a maximum drop density and a maximum rain water content near the top of the cloud.

(ii) *Ground Radar Observations.*—If a cloud producing rain in the manner just described were observed on a radar set, the echo intensity along a vertical section, being proportional to  $\Sigma ND^6$ , would be approximately of the same form

as the rain water content curve of Figure 9. It would therefore be different from that shown in Plate 1, the strongest signal coming not from the falling rain, but from the region where the drops were in suspension.

It is already well known that an intense radar echo is obtained from a horizontal band in many clouds from which rain is falling. It is called the radar "bright band" and has a very characteristic and clearly defined form. In most of the reported instances it has been observed at or just below the freezing level and the explanation(15) which has been advanced for its presence invokes the melting of ice particles or snow-flakes as they fall through freezing level.

Observations made with the radar set referred to in Subsection V (a) show that in addition to the bright band at freezing level, other band structures sometimes occur at heights so far removed from freezing level that they are unlikely

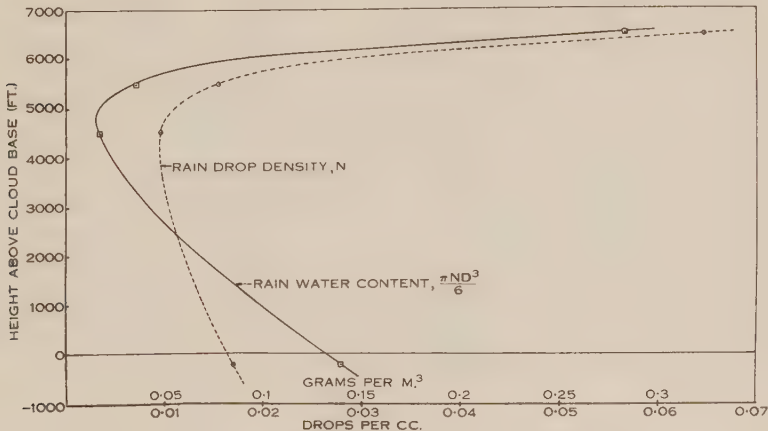


Fig. 10.—Experimental observations of raindrop density and rain water content at different heights in a non-freezing cloud.

to be connected with the melting phenomenon. Two cases which were observed in detail occurred on June 1 and June 6, 1949.

*June 1, 1949.*—On this occasion light showers fell during the morning, the cloud base being at 2000 ft. The cloud tops as determined by aircraft observation were uniform and between 5000 and 6000 ft. When rain was falling, radar echoes were received from a horizontal band overhead at 5000 ft., the band forming and re-forming as the showers passed. It was clearly defined but of low intensity and approximately 1000 ft. thick. A radiosonde record taken at 1800 hours on the same day at Rathmines, some 50 miles north of the point of observation, is given in Figure 12 (a), showing that the freezing level was at 7500 ft. and the temperature in the vicinity of the band from +3 to +5 °C. Furthermore, a distinct inversion existed at 5000 ft. which was probably responsible for limiting the clouds to about this level, and for limiting the raindrop trajectories to a particular height.

The average diameter of the raindrops during a typical shower was measured by allowing them to fall on sensitized paper, and found to be 0.5 mm. Reference to Figure 4 indicates that this agrees closely with the value to be expected from the theory for a radar band 3000 ft. above the cloud base.

*June 6, 1949.*—The conditions were similar to those on June 1 except that drizzle only was falling. The cloud base was at 1500 ft. and the cloud tops estimated to be between 7000 and 8000 ft. A radar band 1000 ft. thick was observed at a height of 3000 ft. fluctuating in intensity with the rainfall. The radiosonde record for that day shown in Figure 12 (b) indicates that freezing level was at 6500 ft. and that the temperature in the vicinity of the radar band was +8 °C. Unlike the previous occasion, there was no evidence of an inversion or change of lapse rate at this height. The average drop diameter was 0.35 mm. and the height of the radar band 1500 ft. above the cloud base, again in approximate agreement with the theory.

(iii) *Airborne Radar Observations of December 1, 1949.*—The conditions already described during the morning of December 1, 1949 persisted in the afternoon and a flight was made about fifty miles inland from the coast to observe any further rain which might occur. The flight took place in an aircraft of the Royal Australian Air Force operated by a special detachment of the Aircraft Research and Development Unit based at Richmond, N.S.W. A temperature sounding made during the ascent gave the curve shown in Figure 12 (c), the probable error in height being 100 ft. and in temperature 1 °C. The cloud base was at 4000 ft. and the cloud tops generally at 10,500 ft. where the temperature was +4 °C. As in the morning, a few cumulus heads pushed up approximately 1000 ft. higher, but at no time did any clouds in the area reach freezing level, which was at 13,000 ft.

TABLE 2  
VARIATION IN HEIGHT OF TOP OF CONVECTIVE CLOUD—  
DECEMBER 1, 1949

Time (hr.)	Height of Top of Cloud (ft.)	Temperature (°C.)
1400	11,000	+4
1419	11,500	+3
1422	12,000	+2
1450	12,500	+1
1520	10,500	+4

At 1400 hours a small head appeared some 500 ft. above the surroundings and it was selected for special observation. It rose steadily during the next hour to a maximum height of 12,500 ft. as indicated in Table 2, after which it collapsed rapidly to the 10,500 ft. level. Flights were made through it at 1419 and 1421 hours at a height of 11,000 ft., just below the top of the cloud. It



would be classed as "very wet" for a cumulus cloud and, as would be expected from the temperature at that level, it contained no ice or snow. The aircraft was fitted with a radar set and during both flights through the cloud a substantial rain echo was observed between the aircraft and the ground. Thereafter flights were made over the top of the cloud at approximately 5 minute intervals, keeping the rain area under observation. The radar set was an SCR 717 similar to that employed for the ground observations, with the antenna mounted so that its axis of rotation was along the fore-and-aft line of the aircraft. As a result the indicator gave a radar cross-section through a plane passing through the aircraft at right angles to the line of flight, instead of the normal plan-position display. Typical photographs of the scan are shown in Plate 2 in which the small bright circle corresponds to the position of the aircraft, the echo pattern immediately below it to radar echoes from the rain, and the bright area below to the echo from the ground. The ground echo appears as an arc of a circle rather than a horizontal line owing to the finite beam width of the antenna system.

A sequence of photographs taken between 1427 and 1504 hours over the region giving the most intense rain echo appears in Plate 2. The lateral extent of the shower was approximately 2 miles in both an east-west and a north-south direction and the photographs show that for the greater part of the time the rain extended from a height of 10,000 ft. to the ground, which was 3000 ft. above sea-level. The echo pattern in the east-west plane showed a distinct shear corresponding to the fact that the winds up to 5000 ft. were generally from the east and above that height from the west. The echo intensity was maintained from the time it was first observed at 1419 until 1450 hours, after which it gradually decreased and disappeared at 1515 hours. The aircraft then descended to the base of the cloud to establish its position and at 1530 hours found only light drizzle between the base of the cloud and the ground. A point of considerable importance in relation to the present theory is that the echo intensity remained strong until 1450 hours, that is, while the cloud was in the process of building up, but fell off in intensity as soon as the cloud had passed its maximum development.

Due to the limited range of echo intensities which can be recorded on a pictorial scan, the photographs in Plate 2 are not suitable for a determination of radar echo intensities. During the flight over the top of the cloud at 1455 hours, therefore, measurements of radar echo intensity against height were made on a separate cathode ray tube which displayed the echo amplitude against height along the line defined by the bright sector in Plate 2 (*d*). These are plotted in Figure 11 as radar echo intensity in decibels above receiver noise level against height in feet, allowance being made for the known variation of echo intensity with distance from the point of observation. Three such measurements were made during a single north-south traverse at 15, 25, and 35 seconds after 1455 hours respectively. All three show a decided maximum in the region between 8500 and 9000 ft. Reference to Figure 12 (*c*) indicates that this corresponds in height to a slight inversion in the temperature lapse rate.

These results are again in qualitative agreement with the theory in showing a region of high water content at a definite height within the cloud.



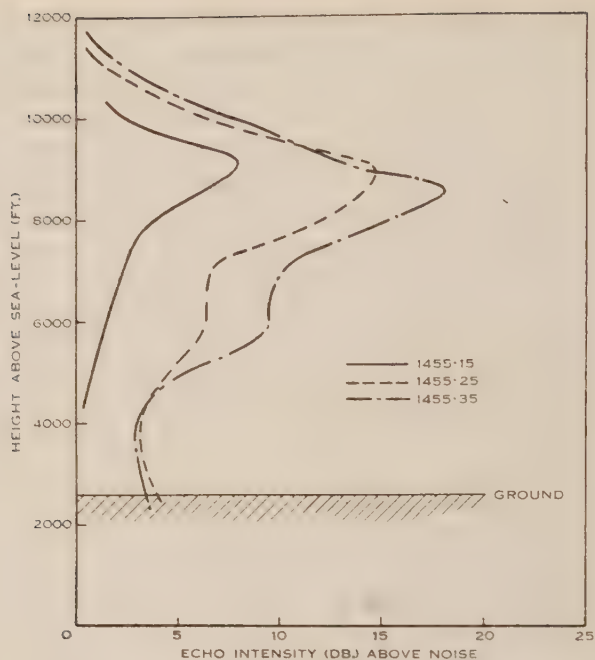


Fig. 11.—Experimental observations of radar echo intensity against height of rain echoes in a non-freezing cloud on December 1, 1949.

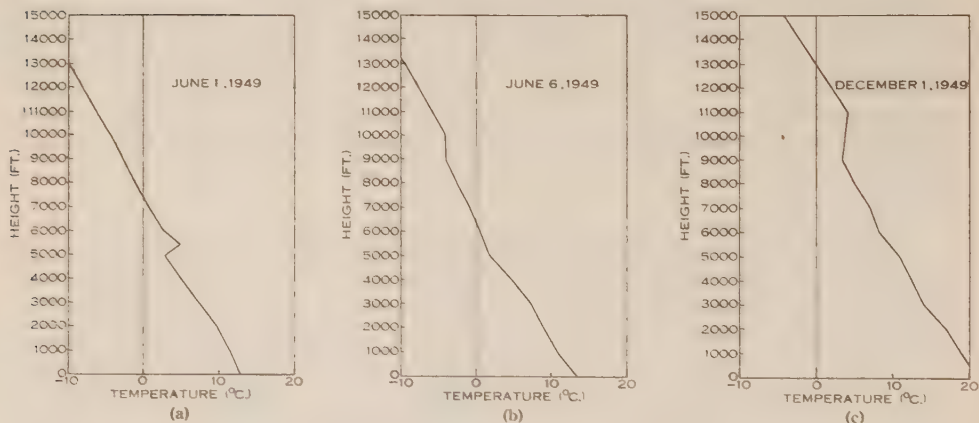
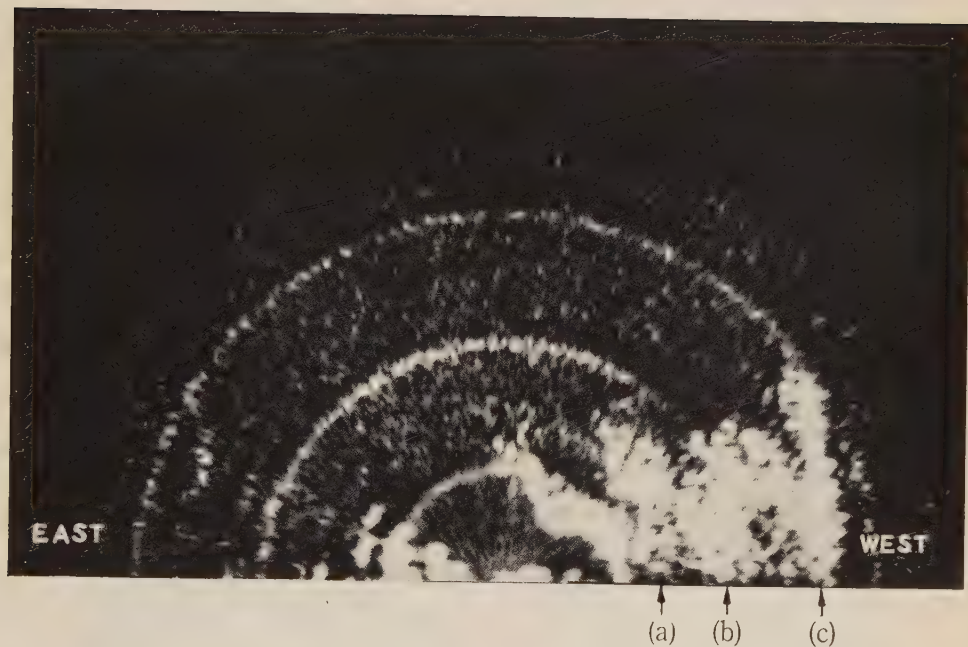


Fig. 12.—The 1800 hour radiosonde records for Rathmines on June 1 and June 6, 1949, and the aircraft sounding for December 1, 1949.

## VI. FACTORS TENDING TO MAKE DROP TRAJECTORIES COINCIDE

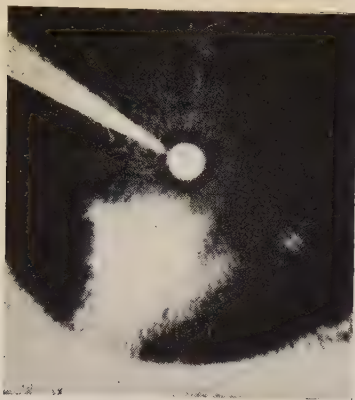
It appears from the foregoing account that the factor most likely to give a high rain water content at a particular height is a decrease or a discontinuity in the vertical air velocity. This may be associated with :

- (i) A temperature inversion or a change of lapse rate in the atmosphere.
- (ii) The top of a convective cell.

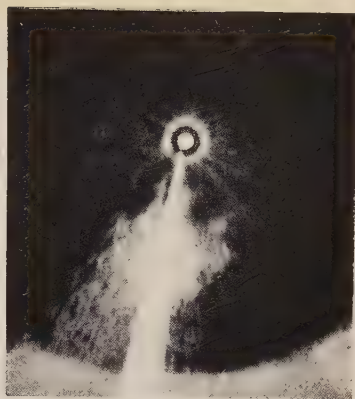


Radar echoes of rain falling from a non-freezing cloud on December 1, 1949.

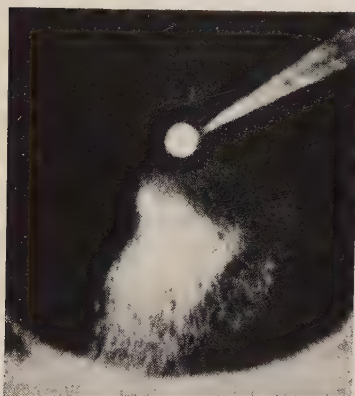




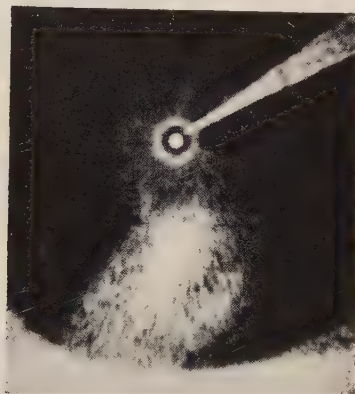
(a) 1427 HOURS



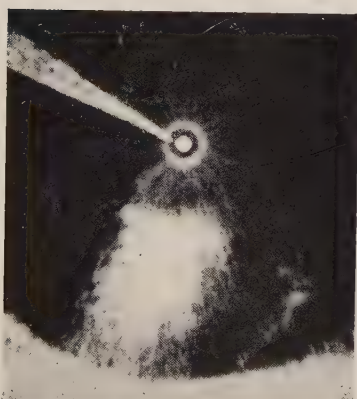
(d) 1455 HOURS



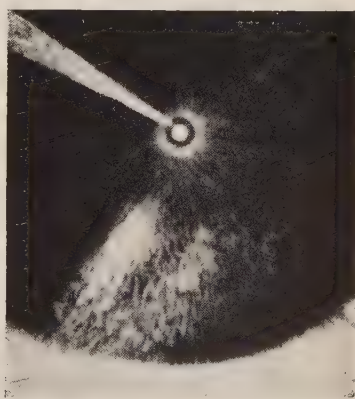
(b) 1446 HOURS



(e) 1500 HOURS



(c) 1450 HOURS



(f) 1504 HOURS

0 5000 10000  
SCALE IN FEET

Airborne radar observations of rain falling from a non-freezing cloud on December 1, 1949. The aircraft height was 14,000 ft. in all cases except (a) when it was 12,500 ft.





The example of December 1, 1949, was of the first kind, while that of June 1, 1948, was of the second kind. On June 1, 1949, the concentration of water appeared both at the top of the cloud and at an inversion, while on June 6, 1949, it appeared to be near the middle of the cloud but there was no evidence for a change of lapse rate at that level. The radiosonde record, however, referred to a different place and a different time and it is quite possible that a discontinuity in the lapse rate was present either locally or of a magnitude too small to be measured.

It might be suggested that the results obtained on June 1, 1948, and June 1, 1949, when the concentration of water drops was observed at the top of the cloud, could be accounted for by the mechanism of radiation cooling postulated by Reynolds(16). On both of these occasions, however, the cloud tops were exposed to the sun and were probably experiencing a net gain of heat by radiation rather than a loss. It is, therefore, unlikely to have been the mechanism at work and, as it could not explain those cases in which the concentration of water drops occurred near the centre of the clouds, the author is led to the conclusion that coalescence was the predominant mechanism in the examples cited in the present paper.

## VII. CONCLUSION

Calculations have been made of the results to be expected if cloud droplets coalesce when they come into collision in natural clouds. The mechanism postulated is one in which the cloud droplets first grow by condensation in their ascent through the cloud, some fraction of their number then growing further by coalescence. Those which grow by coalescence ascend until they can no longer be sustained in the upward air current, after which they fall back through the cloud. They grow still further in their descent to the cloud base and finally emerge as rain. It is found that the mechanism is capable of accounting for raindrops of a wide range of sizes and might therefore be the process at work when rain has been observed to fall from clouds consisting wholly of water droplets.

The maximum height attained by the drops and the final diameter of the raindrops emerging from the base of the cloud are shown to be nearly a linear function of the vertical air velocity. The largest raindrops would therefore be expected to fall from clouds with the greatest convective activity. The time for rain to form by the process should, on the other hand, be inversely proportional to the vertical air velocity, clouds of low activity taking a long time to precipitate. These results are consistent with the properties of natural rain.

Experimental observations of rain falling from non-freezing clouds have distinguished two main types. The first of these appears to be the general case in which, due to turbulence or lack of uniformity within the cloud, the drops have a variety of trajectories. It is characterized by a gradual increase of raindrop diameter or rainfall intensity downward through the cloud. The second corresponds to the case in which the raindrop trajectories tend to coincide, in which case a concentration of relatively large drops would be expected to

form at some height above the cloud base. This has been observed experimentally as a region from which intense radar echoes are received and as a region of high water content during aircraft flights through rain-producing clouds. The many points of agreement between these observations and the deductions from the theory lead to the conclusion that coalescence plays an important part in the formation of rain from non-freezing clouds.

Since the great majority of clouds do not produce rain, it is appropriate to conclude by considering briefly the conditions which need to be met before rain can form by the condensation-coalescence process. Apart from the basic requirement that coalescence shall occur as a result of collisions between cloud droplets, the following five conditions are necessary :

- (1) A distribution of cloud droplet sizes must exist so that the droplets have an opportunity of falling relative to one another and coming into collision. It is probable that this condition is, in fact, met in the great majority of clouds.
- (2) The width of the cloud must be such that drops which grow by coalescence will not be carried out of the cloud by a wind shear or another similar factor. This requirement is common to almost any theory of rain formation.

Coming now to the conditions which arise from the theory as developed in the present paper :

- (3) The cloud droplets must attain a certain minimum size before collisions are frequent enough to give a reasonable number of raindrops. The present state of knowledge on the collision process is not good enough to define this limit at all accurately.
- (4) The vertical air current in a cloud must be maintained long enough for the growth process to be completed. It is a matter of observation that many convective clouds do not meet this requirement, going through their whole cycle of growth and dissipation in a shorter time than that required for raindrops to form by coalescence.
- (5) For a given upward air velocity the depth of the cloud must be greater than that required for the drops which grow by coalescence to come into equilibrium in the upward air current.

If any of these conditions are not met, then rain is unlikely to form by the condensation-coalescence process.

#### VIII. ACKNOWLEDGMENTS

The author is grateful to Mr. E. J. Smith, Division of Radiophysics, C.S.I.R.O., for permission to include his experimental observations of drop density and rain water content, and to Mr. J. Warner for assistance with the radar observations. Grateful acknowledgment is also made to the pilots and crew of the special detachment of A.R.D.U. which made the flight observations possible.

## IX. REFERENCES

- (1) BERGERON, T.—*P.V. Mét. Un. Géod. Géophys. Int.* Lisbon. pp. 156-78 (1933).
- (2) FINDEISEN, W.—*Met. Z.* **56** : 365-8 (1939).
- (3) HOUGHTON, H. G.—*Bull. Amer. Met. Soc.* **19** : 152-9 (1938).
- (4) WEXLER, H.—*Bull. Amer. Met. Soc.* **26** : 156 (1945).
- (5) KOTSCH, W. J.—*Bull. Amer. Met. Soc.* **28** : 87-9 (1947).
- (6) HUNT, T. L.—*Met. Mag.* **78** : 26 (1949).
- (7) FRÖSSLING, N.—*Gerlands Beitr. Geophys.* **52** : 170-216 (1938).
- (8) HOUGHTON, H. G.—*Physics* **4** : 419 (1933).
- (9) RAYLEIGH, LORD.—*Sci. Pap.* **1** : 372-6 ; **2** : 103-15. (Cambridge Univ. Press, 1900.)
- (10) LANGMUIR, I.—*J. Met.* **5** : 175-92 (1948).
- (11) LAWS, J. O.—*Trans. Amer. Geophys. Un.* **22** : 709-21 (1941).
- (12) HOWELL, W. E.—*J. Met.* **6** : 134-49 (1949).
- (13) BYERS, H. R., and BRAHAM, R. R.—*J. Met.* **5** : 71-86 (1948).
- (14) SMITH, E. J.—*Quart. J.R. Met. Soc.* (Communicated.)
- (15) BYERS, H. R., and COONS, R. D.—*J. Met.* **4** : 75-81 (1947).
- (16) REYNOLDS, O.—*Collective Works* **1** : 214. (Cambridge Univ. Press, 1900.)



# OBSERVATION OF PRECIPITATION WITH AN AIRBORNE RADAR

By E. J. SMITH\*

[*Manuscript received December 19, 1949*]

## *Summary*

A radar set with its beam directed downwards has been mounted in an aircraft and used for observation of natural and artificially induced precipitation. It provides information on where the precipitation originates within a cloud and the way in which it grows. A measure is obtained of the area and duration of precipitation reaching the ground, and an estimate of the precipitation rate.

As an example of the use of this equipment, an account is given of a typical experiment in which precipitation was artificially induced by seeding a cloud with dry ice.

## I. INTRODUCTION

Studies of the formation of precipitation, both natural and artificial, have been seriously hampered by the lack of instruments capable of measuring many of the quantities involved. Microwave radar forms a very valuable addition to the available range of measuring instruments, and it is now being used to a considerable extent in meteorological research.

Its utility depends on the fact that radio waves are reflected from water drops and ice particles in the atmosphere, the intensity of the echo from a given type of particle being proportional to the number of particles present per unit volume of the atmosphere, and to the sixth power of their diameter. The echo intensity does not, unfortunately, enable drop size and concentration to be determined uniquely, but if a drop-size distribution is assumed, an estimate can be made of drop concentration and precipitation rate.

Most of the radar observations of precipitation which have been reported to date have been made with ground installations. A disadvantage of observing from a fixed site is that it is not always in the right place with respect to the phenomenon which has to be investigated. This is particularly true of experiments on the artificial production of rain which may be performed some hundreds of miles from the operating base and it is preferable in these cases to use an airborne radar. This paper describes a simple installation of this kind, gives the results of a typical experiment, and the method of reduction of the results.

## II. DESCRIPTION OF EQUIPMENT

The radar is a modified version of the SCR717, operating at a wavelength of 10 centimetres. It is installed in a Dakota aircraft of the R.A.A.F. with a fixed antenna pointing vertically downwards. The antenna dish is 29 inches

\* Division of Radiophysics, C.S.I.R.O., University Grounds, Sydney.

long and 18 inches wide giving a beamwidth, to half power, of approximately 10 degrees along the fuselage of the aircraft and 16 degrees across it; the pulse duration is 1.2 microseconds. A photograph of the antenna installation is given in Plate 1, Figure 1.

A special display unit is used consisting of two cathode-ray tubes mounted one above the other with a watch face between. The upper tube is intensity-modulated with a time base which reads from 0 to 20,000 feet, or 0 to 30,000 feet, range from the aircraft increasing downwards. The lower tube carries the same signal but is amplitude-modulated. The traces are recorded photographically on continuously moving 35 mm. film, a typical record being shown in Plate 1, Figure 2. A continuous record is obtained of the top tube while a series of fixed pictures of the bottom tube and the watch is obtained by flashing a light once every five seconds. In the upper trace of Plate 1, Figure 2, the ordinates represent height in feet, and the abscissae (reading from right to left) represent time or distance flown at a constant air speed. The bright line *A* represents the path of the aircraft and the line *B* is the echo from the ground, the distance between them indicating the height of the aircraft. The bright mass *C* is an echo from precipitation between the aircraft and the ground. On the lower trace a succession of pictures is obtained giving the intensity of the precipitation echo. On a typical trace, *D* represents the position of the aircraft, *E* the precipitation, and *F* the ground.

The particular record shown in Plate 1, Figure 2, was obtained during an experiment in which dry ice was dropped into a cumulus cloud extending from 6000 to 17,000 feet. The aircraft was flying at 18,000 feet, the height of the ground being 2500 feet. Precipitation echoes were first observed 11 minutes after the dry ice was dropped, and those shown in Plate 1, Figure 2, appeared 5 minutes later.

### III. INTERPRETATION OF RADAR ECHOES FROM PRECIPITATION

The characteristics of radar echoes from different types of precipitation have been discussed by Ryde(1). Assuming the drop-size distribution given by Laws and Parsons(2), Ryde has computed the equivalent echoing cross-sectional area for rain of given intensity as observed on a radar set of given characteristics.

The relation between echo intensity, range, and echoing area for the radar set under discussion has been determined by observation of echoes from metallized spheres, a sphere of diameter 2 feet being visible for 3 miles. This has been confirmed by measurements of transmitter power and receiver sensitivity. Periodical checks of the overall sensitivity showed a variation of less than 3 db. From these data and assumptions it has been possible to calculate the precipitation rate corresponding to any rain echo. A set of curves relating the precipitation rate, the radar echo intensity, and range obtained in this way is given in Figure 1.

In the case of ice crystals and snow-flakes, no relation has been established between precipitation rate and size of the crystals or snow-flakes, and precipitation rate can only be calculated from radar-echo intensity if the size of the snow-flakes is known.

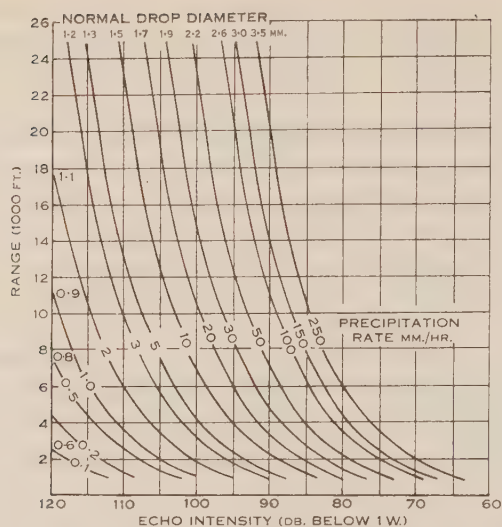


Fig. 1.—Relation between signal strength, precipitation rate, and range for rain.

Some of the properties of snow-flakes and ice crystals have been given by Hooper and Kippax(3) and those relevant to the present investigation are summarized in Table 1.

TABLE 1  
RELEVANT PROPERTIES OF SNOW-FLAKES AND ICE CRYSTALS

				Snow-Flakes			Ice Crystals	
Diameter (mm.)	..	..	..	20	10	5	2	1 (extra- polated)
Mass (mg.)	..	..	..	210	26	3	0.025	0.006
Terminal velocity (m./sec.)	..	..	..	1.4	1.2	1	0.4	0.4
Equivalent raindrop diameter (mm.)				7.5	3.7	1.8	0.37	0.22

From the quantities in this table the relations between echo intensity, range, and precipitation rate can be calculated from Ryde's results for different sizes of snow-flakes and these are given in Figures 2-6.

#### IV. DESCRIPTION OF A TYPICAL EXPERIMENT

A description of a typical experiment is given as an example of the use of the downward-looking radar. On May 5, 1949, a shower was artificially induced by dropping dry ice into a cumulus cloud. Over the coastal range, 100 to 150 miles west of Sydney, there was four-eighths cumulus cloud with tops at 11,000 to 12,000 feet at temperatures of  $-11$  to  $-12\frac{1}{2}$  °C., and with bases

at 6000 feet at a temperature of  $-2^{\circ}\text{C}$ . The winds were westerly at all heights up to 10,000 feet with a wind speed of 20 m.p.h. and with no evidence of a wind shear above.

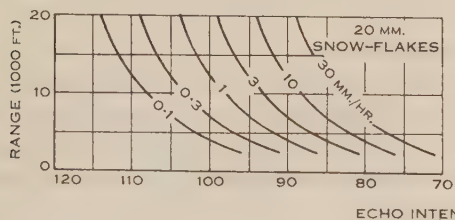


Fig. 2

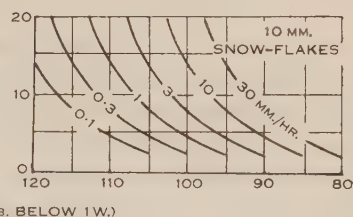


Fig. 3

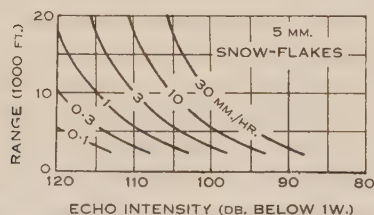


Fig. 4

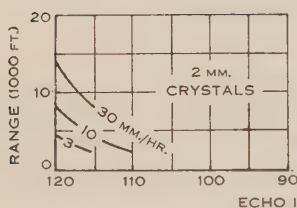


Fig. 5

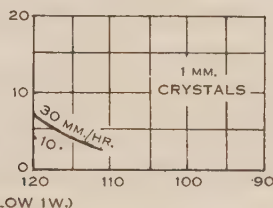


Fig. 6

Fig. 2.—Relation between signal strength, precipitation rate, and range for 2 cm. snow-flakes.

Fig. 3.—Relation between signal strength, precipitation rate, and range for 1 cm. snow-flakes.

Fig. 4.—Relation between signal strength, precipitation rate, and range for 5 mm. snow-flakes.

Fig. 5.—Relation between signal strength, precipitation rate, and range for 2 mm. ice crystals.

Fig. 6.—Relation between signal strength, precipitation rate, and range for 1 mm. ice crystals.

One of the larger clouds situated about 30 miles north of Bathurst was selected for an experiment. The base covered an area of five by three miles, aligned NNW.-SSE.; the top was at 12,000 feet,  $-12\frac{1}{2}^{\circ}\text{C}$ . and had the characteristic well-rounded appearance of clouds composed of water drops. No radar echoes could be obtained from this or neighbouring clouds. At 13.47 hours the aircraft flew over the top of the cloud along its major dimension, at a



height of 12,300 feet. Eighty pounds of granulated dry ice was dropped at a rate of 40 pounds per mile in the manner described previously by Squires and Smith(4). The aircraft then climbed to 13,500 feet and flew backwards and forwards over the top of the cloud while radar observations were made.

(a) *Visual Effects Observed*

The sequence of events observed visually from the aircraft was similar in the main characteristics to those which have previously been described by the author(5). Thirty minutes after the dry ice was dropped the top of the cloud lost its characteristic appearance of water drops and changed to that associated with ice crystals. After 49 minutes the aircraft descended to the bottom of the cloud from which snow was falling over an area about 3 miles by 1 mile aligned NNW.-SSE. At this time, the base of the cloud had become indistinct and was merged with the snow, the cloud having turned completely to ice crystals and snow. There was no indication of any effects having spread to neighbouring clouds, the bases of which remained clearly defined. After 52 minutes, the aircraft flew at 5000 feet through the precipitation which consisted of ice crystals or small snow-flakes whose diameter was estimated as 1 mm., visibility through the snow being approximately  $\frac{1}{2}$  mile. At 4000 feet the snow was melting but the aircraft could not descend to investigate conditions below this height owing to the presence of mountain peaks. The aircraft left the area after 55 minutes and the cloud was last seen after 60 minutes, by which time it had almost dispersed and the precipitation had stopped.

(b) *The Radar Observations*

Returning now to consideration of the radar observations which were made, these are summarized in Table 2, which gives the times at which the aircraft crossed the top of the cloud, the aircraft heading, the maximum echo intensities,

TABLE 2  
SUMMARY OF RADAR OBSERVATIONS

Time	Course	Maximum Intensity (db. below 1 watt)	Precipitation Rate at Ground (mm./hr.)	Remarks
6 Min. before ..	E.	—	—	Preliminary observation
Zero (13.47 hr.) ..	SSE.	—	—	Dry ice drop
5 Min. after .. ..	NNW.	—	—	
8 Min. after .. ..	NE.	—	—	
11 Min. after .. ..	SSW.	120	—	Aircraft track slightly displaced to west
13 Min. after .. ..	NNW.	114	1.5	
18 Min. after .. ..	NW.	112	2.7	
27 Min. after .. ..	NNW.	114	2.3	
33 Min. after .. ..	NE.	116	2.0	
38 Min. after .. ..	SW.	—	—	
42 Min. after .. ..	N.	—	—	

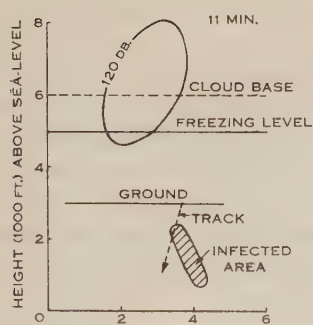


Fig. 7

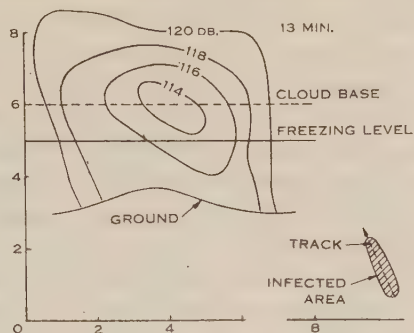


Fig. 8

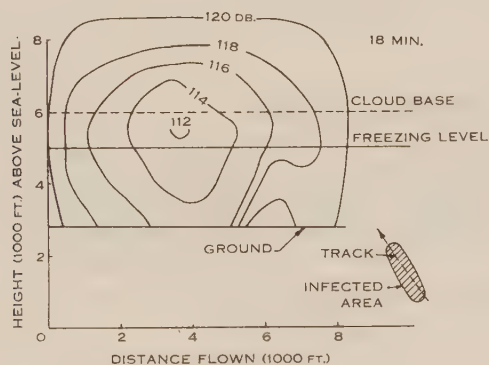


Fig. 9

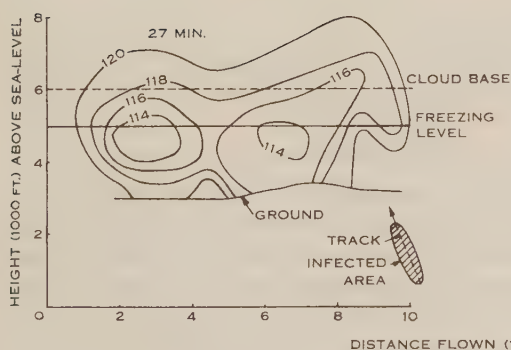


Fig. 10

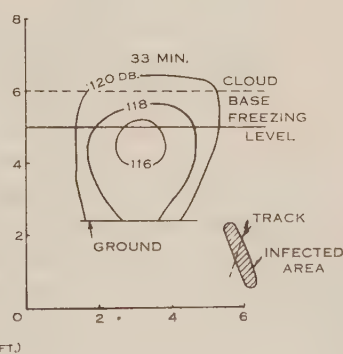


Fig. 11

Figs. 7-11.—Experiment on May 5, 1949. Plot of echo intensity against height and distance flown. Fig. 7.—Eleven minutes after dry ice drop. Fig. 8.—Thirteen minutes after dry ice drop. Fig. 9.—Eighteen minutes after dry ice drop. Fig. 10.—Twenty-seven minutes after dry ice drop. Fig. 11.—Thirty-three minutes after dry ice drop.

and the precipitation rates, at the ground, deduced from the radar echoes. All the times given are relative to 13.47 hours, the time at which the dry ice was dropped.

In addition flights were made over the edges of the cloud to establish the horizontal limits of the precipitation echoes. They were at all times confined within an area of 3 miles by 1 mile under the cloud which was drifting with the prevailing 20 m.p.h. westerly wind. The intensity of the echoes received is displayed in Figures 7-11 as contours of equal signal strength, the ordinates being height above sea-level and the abscissae, distance flown. Inserted in the bottom right-hand corner of each figure is a diagram representing the infected area and the track of the aircraft relative to it.

Curves of the variation of signal strength with height are shown in Figure 12. They represent vertical sections through the region of maximum echo intensity on the track of the aircraft.

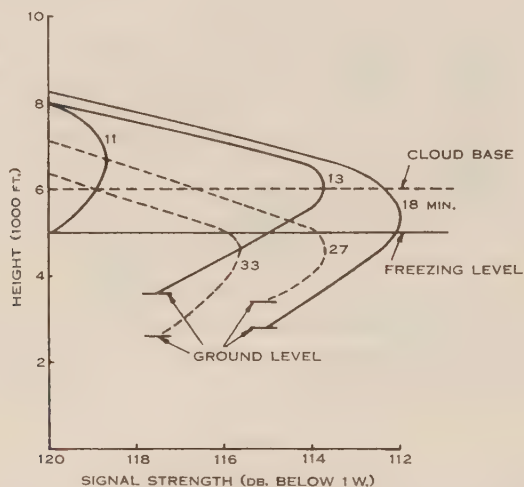


Fig. 12.—Variation of maximum echo intensity with height for various times after dry ice drop.

### (c) Interpretation of Radar Echoes

The radar echoes show that precipitation started about 11 minutes after the dry ice was dropped, growing to a maximum after 18 minutes and then gradually dying away. From previous experience(4), and from visual observation of the form in which the precipitation left the cloud, it is reasonable to assume that the whole of the radar echoes obtained from heights above the freezing level were due to snow-flakes and ice crystals. These will have melted below the freezing level, and echoes originating from below about 4000 feet are assumed to have come from rain. These rain echoes decrease in intensity with decreasing height owing to the increasing range from the aircraft. The rate of decrease of intensity indicates that the precipitation rate was independent of height, and that the rain fell without appreciable evaporation. The rate of precipitation reaching the ground is plotted against time in Figure 13. This is

deduced from the most intense echoes observed; the precipitation rate was somewhat less towards the edges. The minimum rate of precipitation which the radar can detect at that range is 1.4 mm./hr. and this rate was exceeded for 23 minutes, the maximum rate being 2.8 mm./hr. after 20 minutes.

Extrapolating to zero, the precipitation commenced after about 11 minutes and the total duration of rainfall was 50 minutes. It can be calculated from these figures that about 2000 tons of water reached the ground.

Some deductions can also be drawn about the precipitation above freezing level but these are less definite than those concerning the rain owing to lack of information on the particle size. For example, a signal strength of 115 db. below 1 watt, at a height of 7000 feet, as seen 18 minutes after the dry ice drop, would indicate a precipitation rate of 20 mm./hr. if it were composed of 2 mm.

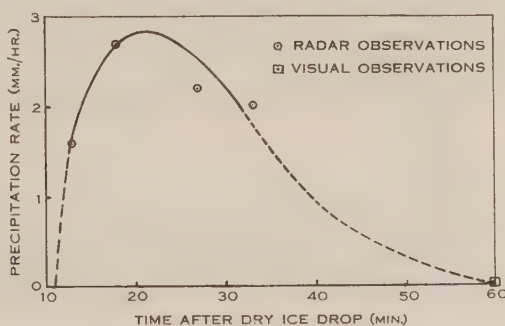


Fig. 13.—Variation of precipitation rate with time at ground level.

ice crystals, or a rate of 0.3 mm./hr. if it were 0.5 cm. snow-flakes (see Figs. 4 and 5). As mentioned earlier the size of the snow-flakes and precipitation rate cannot be uniquely determined from the radar observations. However, an estimate of both can be obtained if it is assumed that there is a degree of uniformity in the precipitation rate across the freezing level. In Figure 12 the contours of radar intensity 18 minutes after the dry ice was dropped indicate a precipitation rate below the freezing level of about 2.7 mm./hr. If the snowfall just above the freezing level had the same precipitation rate then by interpolation between Figures 4 and 5 the diameter of the snow-flakes must have been about 0.4 cm. An independent estimate of the size of snow-flakes may be obtained from the fact that a precipitation rate of 2.7 mm./hr. corresponds to raindrops of 1 mm. in diameter. By interpolation in Table 1 the diameter of each corresponding snow-flake is about 0.4 cm.

The echoes, when first observed 11 minutes after the dry ice drop, extended from 5000 to 8000 feet. The ice crystals which caused the echoes may have formed at higher levels, and grown while falling to the heights at which they were observed. The rate of fall of small ice crystals is given by Hooper and Kippax as 0.4 metre per second, which indicates that they would have formed about 1000 feet above the height at which they were first observed, that is, between 6000 and 9000 feet.



Additional information on the rate of fall of the disturbance is obtained if the height from which echoes of maximum intensity were received is plotted against time, as in Figure 14. The points indicate that the disturbance was initially centred at about 7500 feet, and fell gradually to 4600 feet with an average rate of fall of 2 feet per second.

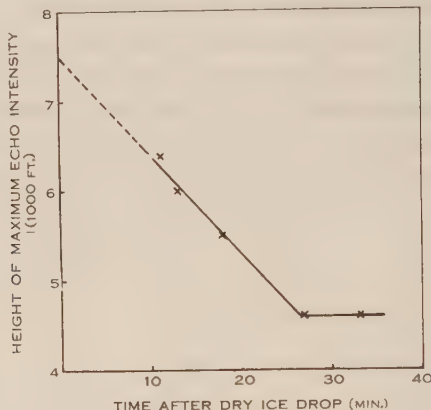


Fig. 14.—Variation of height of maximum echo intensity with time.

#### (d) Discussion of Radar Echoes

Bergeron(6) has described how rain can form in a cloud of supercooled water droplets in which ice crystals form or are introduced. Water evaporates from the droplets and sublines on to the ice crystals, which grow and fall out. Squires and Smith(4) have described a series of experiments in the artificial formation of rain by seeding clouds with dry ice. In that paper (written prior to the introduction of downward-looking radar), deductive evidence was given that the rain so produced was formed in the manner postulated by Bergeron, growing initially from ice crystals. Evidence was also given that even in the case of clouds extending both above and below the freezing level, the Bergeron process must have operated in the part which was colder than freezing.

When experiments are observed by downward-looking radar, as was the case on May 5, 1949, the height at which the disturbance commences is measured directly by means of the radar echoes produced. This represents a valuable advance, and further experiments of the same nature should provide definite information on the processes at work. On May 5 the ice crystals appear to have formed initially at about 7500 feet, which was nearer the bottom than the top of the cloud, at a temperature of about  $-6^{\circ}\text{C}$ . This is surprising, since the top of the cloud was at  $-12\frac{1}{2}^{\circ}\text{C}$ . where crystals would be expected to grow faster, since the vapour pressure difference between water and ice is greater.

Many observations have been made of radar echoes from rain in which there was a marked increase in intensity in a restricted band of height just below the freezing level. The fact that the maximum echo intensity on the present

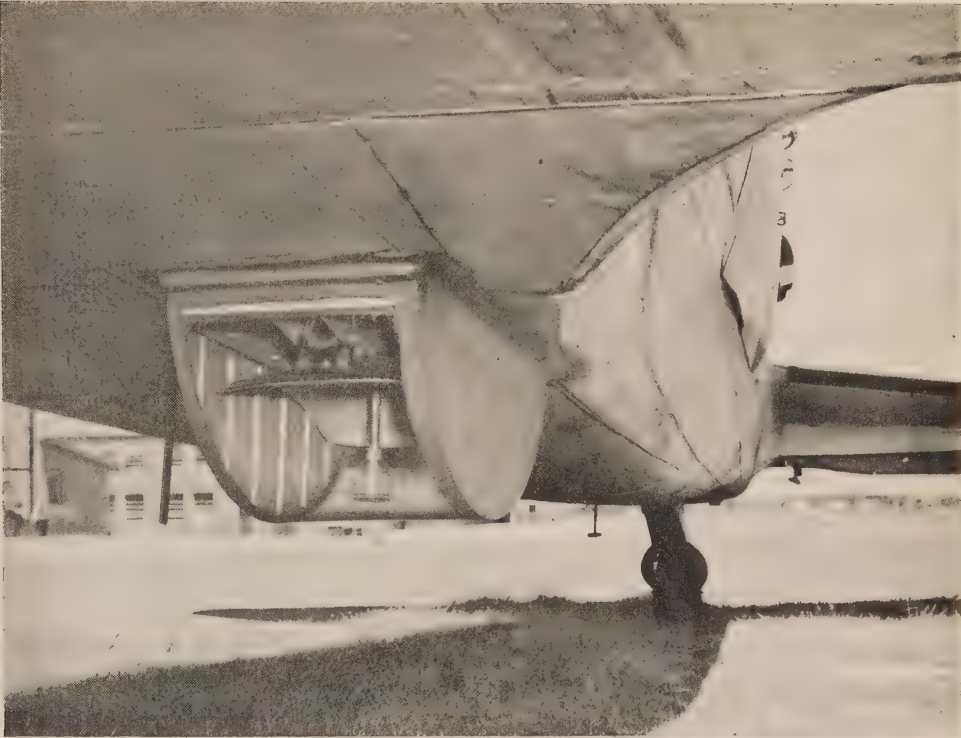


Fig. 1.—Radar antenna under aircraft fuselage.

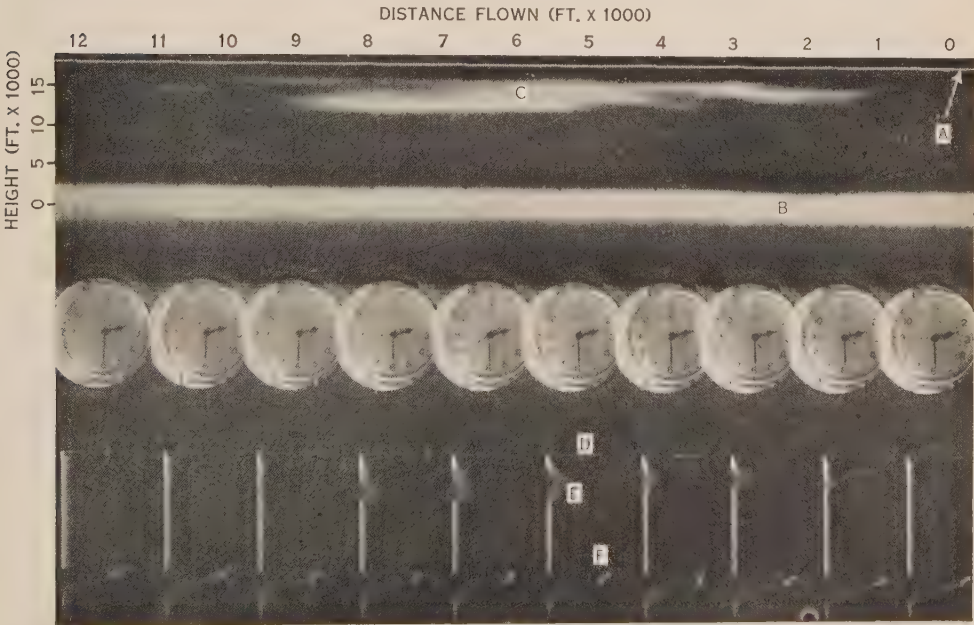


Fig. 2.—Sample photograph of radar display.



occasion was received from a height which fell to 500 feet below the freezing level, and remained there, may be related to this effect.

#### V. ACKNOWLEDGMENTS

Grateful acknowledgment is made to Dr. E. G. Bowen, who suggested the downward-looking radar, to Mr. H. Edwardes, who modified, assembled, and installed the radar set, to Mr. S. Pett, who made the display, to Mr. K. J. Heffernan, who operated the equipment in flight, and to a special detachment of the Aircraft Research and Development Establishment of the R.A.A.F., who provided and flew the aircraft.

#### VI. REFERENCES

- (1) RYDE, J. W.—Report of joint meeting Physical and Royal Meteorological Societies, p. 169 (1947).
- (2) LAWS, J. O., and PARSONS, D. A.—*Trans. Amer. Geophys. Un.*, Pt. II, p. 452 (1943).
- (3) HOOPER, J. E. N., and KIPPAX, A. A.—T.R.E. Rep. No. T2082 (1947).
- (4) SQUIRES, P., and SMITH, E. J.—*Aust. J. Sci. Res. A* **2**: 232 (1949).
- (5) SMITH, E. J.—*Aust. J. Sci. Res. A* **2**: 78 (1949).
- (6) BERGERON, T.—*P.V. Mët. Un. Geol. Geophys. Int.* Lisbon (1933).



# SURVEY OF GALACTIC RADIO-NOISE AT 200 Mc/s.

By C. W. ALLEN\* and C. S. GUM\*

[*Manuscript received December 5, 1949*]

## *Summary*

A survey of galactic radio-noise at 200 Mc/s. has been made for the region of the sky south of declination  $45^\circ\text{N}$ . Results are expressed graphically in watts  $\text{m}^{-2} (\text{c/s.})^{-1}$  (steradian) $^{-1}$ , the maximum value being  $20 \times 10^{-21} \text{ W. m}^{-2} (\text{c/s.})^{-1}$  (steradian) $^{-1}$  at galactic latitude  $-3^\circ$  and longitude  $322^\circ$ . Corrections have been made for the aerial beam width. No measurement has been made of the residual intensity from the coldest parts of the sky (the galactic poles) and this quantity, when known, should be added to the results quoted. The galactic noise provides a permanent and convenient standard for solar noise measurements, and its intensity for this purpose is evaluated.

## I. INTRODUCTION

One of the advantages that radio-astronomy has over optical astronomy is that radio waves can penetrate more deeply through interstellar space near the centre of the galaxy. Radio-noise intensity measurements can thus give information on the nature of interstellar matter and the structure of the galaxy. Accurate surveys at various wavelengths of the radio-noise intensity from the whole sky form an important observational requirement for advances of this character.

The observations need to be expressed in absolute units, and a correction has to be made for the aerial pattern. Sensitivity, aerial beam width and the method of correcting for it, and horizon effects, are liable to vary from one observer to another and the safest way to get reliable results is to obtain independent observations with a variety of conditions.

The radio-noise set at the Commonwealth Observatory was prepared by the Division of Radiophysics, C.S.I.R.O., Sydney, essentially for the observation of solar-noise(1). It is also adaptable for galactic noise although the aerial beam width is somewhat wider than desirable for that purpose. The following reasons have prompted the survey of galactic noise described in this paper: (a) The desirability of having independent estimates of galactic noise from a number of observers, (b) the need for observations from a position where the galactic centre may be observed near the zenith, and (c) to have both solar and galactic noise observations made with the same set, so that intensities can be compared on the same scale. Slow variation of quiet day solar noise can be detected by frequent measurements of the sun/sky noise intensity ratio.

## II. EQUIPMENT AND AERIAL PATTERN

The aerial consists of an array of four Yagis at the corners of a 160 cm. square. Each Yagi has four directors and one reflector. The aerial is mounted

\* Commonwealth Observatory, Mt. Stromlo, Canberra, A.C.T.

on a polar axis which allows continuous movement in right ascension, and declination movement in steps of  $10^\circ$ , set at  $\pm 5^\circ$ ,  $\pm 15^\circ$ , etc. Energy from the aerial is fed to a high sensitivity superheterodyne receiver.

For solar observations the receiver output is connected to a recording milliammeter, but for galactic noise greater sensitivity is required and the readings have been made on a backed-off microammeter.

The aerial sensitivity pattern has been determined by observing the output variation as the aerial was moved through the sun in both E.-W. and N.-S. directions. The two patterns are almost identical and the mean is shown in Figure 1. These measurements were made when the solar noise was steady, and corrections for galactic background noise have been made where necessary. The mean curve was adopted and assumed to apply for all directions. This assumption is not precisely accurate since the Yagi array is in the form of a square, but considerations of the interference pattern show that no appreciable error is introduced by assuming circular symmetry. We have eliminated the

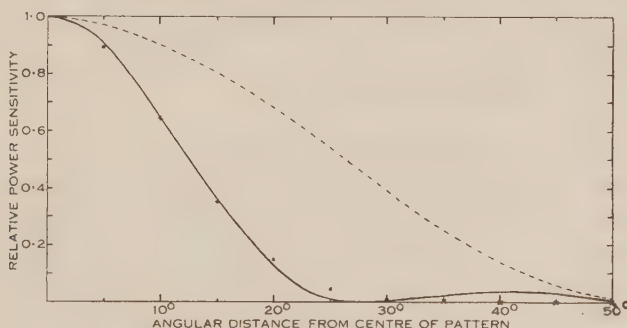


Fig. 1.—Aerial sensitivity pattern.

— Mean curve used for analysis.

- - - Single Yagi.

x Calculated values at  $45^\circ$  to N.-S. direction.

effect of interference to determine the aerial sensitivity for a single Yagi. This is also shown in Figure 1. The effect of interference at  $45^\circ$  to N.-S. direction may be calculated and the result is seen in Figure 1 to be very similar to the curve adopted.

It may be noticed that the aerial pattern is very nearly Gaussian. The Gaussian character of the distribution has an advantage which we use later, that the distribution at right angles to a linear source when observed with such an aerial is the same as the distribution for a point source. This enables us in Section VI to treat the correction for aerial pattern in the first approximation as a one-dimensional problem.

### III. OBSERVATIONS

The observations were made at the Commonwealth Observatory (long. 9 hr. 56 min. E.; lat.  $35^\circ 20'$  S.) during January, April, May, and June 1949 at times when the solar activity was low.

The bulk of the observations consisted of right ascension sweeps in half-hourly steps at fixed declinations from  $45^{\circ}\text{N.}$  to  $85^{\circ}\text{S.}$  The sweeps covered at least six hours of right ascension at all declinations and 12 hours for some southern ones. Normally a set of observations consisted of three sweeps in each direction made as rapidly as possible in order to reduce the effect of zero drift. The right ascension sweeps were interconnected by meridian sweeps which went from  $45^{\circ}\text{N.}$  to  $15^{\circ}$  below the south pole. In order to cover the whole sky, observations were made at various times of the day and night, and since they extended over several months the effect of solar noise could be eliminated. About 15 such sets of observations were made at each declination.

Throughout the observations there was no absolute standard of zero intensity, only differences of intensity were measured. In the final results the lowest measured intensity is taken as zero, and within the limits of accuracy this minimum is the same over large areas of the sky near the galactic poles. Unfortunately it is not possible with the present equipment to measure the residual radiation coming from these areas. Bolton (personal communication) has extrapolated his own measurements to 200 Mc/s. and gives  $100^{\circ}\text{K.}$  as a probable temperature for the coldest parts of the sky, or the minimum intensity  $1.2 \times 10^{-21} \text{ W. m.}^{-2} (\text{c/s.})^{-1} (\text{steradian})^{-1}$ .

#### IV. REDUCTION OF OBSERVATIONS

The method of assembling the results was to plot all the sets of observations for a given declination against right ascension. Each set of observations then gave a curve which had an arbitrary zero of intensity. The various curves were then displaced vertically until the best fit was obtained in the low intensity parts. Considerable overlapping in right ascension of these curves made it possible to obtain a mean curve showing the variation of galactic noise for the complete 24 hours of right ascension.

It was found, however, that the readings were affected by (a) proximity to the horizon, and (b) proximity of the aerial parts to the southern polar axis pillar. Empirical corrections for both of these effects were made as follows.

In the preliminary right ascension plots for the declinations  $-5^{\circ}$ ,  $-15^{\circ}$ , and  $-25^{\circ}$  the values within two hours of the horizon were not used for fitting one set of observations to another. After a smooth curve was drawn the differences between measured values near the horizon and the smoothed values were plotted as a function of altitude. The plots are shown in Figure 2 and the mean curves have been used for correcting all observations within  $30^{\circ}$  of the horizon. The direct use of Figure 2, however, had to be modified when areas of high intensity were close to the horizon. In such cases there was loss of power due to the fact that part of the source was below the horizon. Approximate corrections were made, based on the fraction of the source obscured by the horizon, but these were rather uncertain and observations of the intense areas when near the horizon were given low weight.

Whenever the aerial came close to the southern polar axis pillar an increase of the microammeter reading could be detected. Some correction for this effect was needed for the meridian sweeps which extended to  $15^{\circ}$  below the pole.

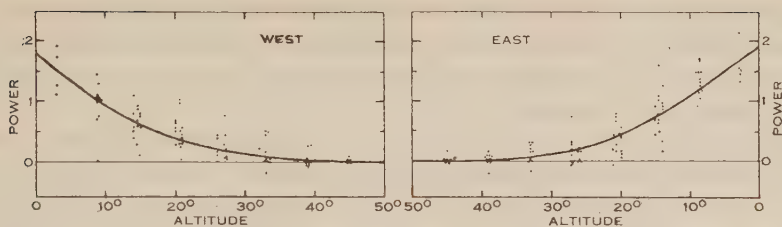


Fig. 2.—Horizon correction. Unit of power  $10^{-22}$  watts  $\text{m}^{-2}$   $(\text{c/s.})^{-1}$ .

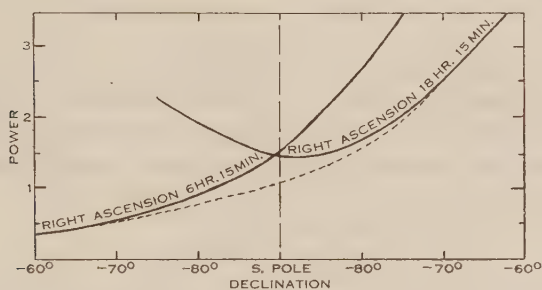


Fig. 3.—Spurious increase of power at the pole. Unit of power  $10^{-22}$  watts  $\text{m}^{-2}$   $(\text{c/s.})^{-1}$ .

—— Measured power along line of constant right ascension.  
 ---- Adopted galactic power.

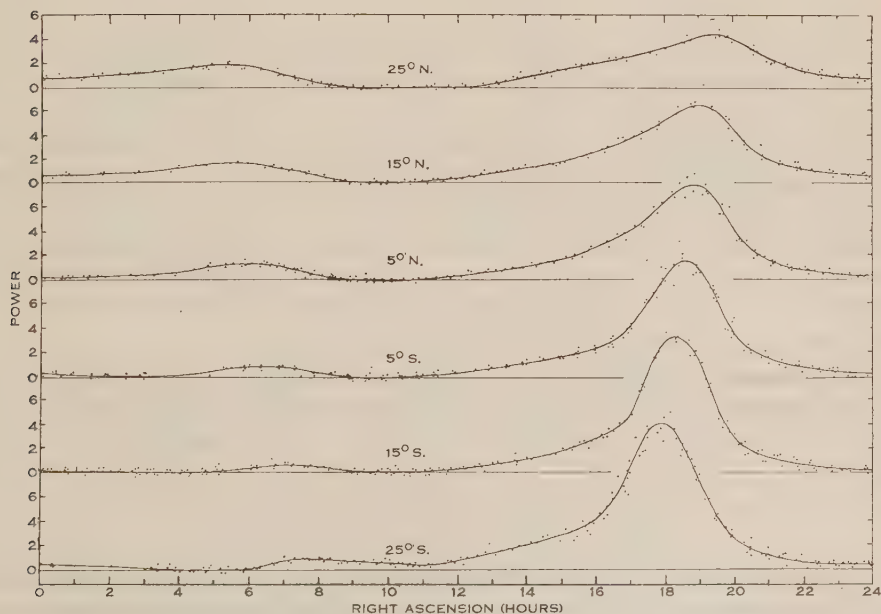


Fig. 4.—Galactic power at various declinations. Unit of power  $10^{-22}$  watts  $\text{m}^{-2}$   $(\text{c/s.})^{-1}$ .



The magnitude of the effect could be determined by plotting meridian observations separated by 12 sidereal hours. An example is shown in Figure 3. There is very little ambiguity possible in filling in the true curve, and then differences are used to give a correction.

After making these corrections the right ascension sweeps were replotted and the intensity levels adjusted to fit the meridian sweeps. Figure 4 shows the scatter of the individual observations for declinations 25°N. to 25°S. The mean curves in this diagram are immediately available for standardizing the solar-noise intensity as explained in Section VII.

All intensities in this paper refer to the total intensity on both components of polarization.

## V. CALIBRATION

The units used for both observation and analysis were the reading differences on the output microammeter. In order to convert these into absolute units a procedure outlined by Pawsey and McCready(2) was followed. Suppose  $T_A$  to be the effective temperature of the sky over the acceptance cone of the aerial, then the power output from the set will be

$$P = kgB[T_A + (N-1)T_R],$$

where  $k$  is the Boltzmann constant of  $1.38 \times 10^{-23}$  joule deg.  $^{-1}$ ,  $g$  the receiver gain,  $B$  the integrated noise acceptance bandwidth,  $N$  the noise factor of the set, and  $T_R$  the ambient temperature taken as 300°K. If we regard the coldest part of the sky as being at zero temperature, and consider the change  $\Delta P$  in turning to a region at  $T_A$  then

$$\frac{\Delta P}{P} = \frac{L_1 L_2 T_A}{(N-1)T_R}, \quad \dots \dots \dots (1)$$

where we have now introduced loss factors  $L_1$  and  $L_2$ .  $L_1$  represents the loss in the aerial transmission line and in the present case is estimated to be 0.65. The factor  $L_2$  represents the loss from other causes. This has varied from time to time. Assuming that at maximum sensitivity  $L_2$  is 1.0 (as in January 1949) the factor  $L_2$  for most of the present observations (April, May, and June 1949) was 0.67. The noise factor  $N$ , defined as the receiver noise power relative to that delivered by a resistor at ambient temperature  $T_R$ , was determined by injecting an artificial noise source from the diode noise generator. During the observations  $N=5.2$ . The quantity  $\Delta P/P$  may be determined from the microammeter deflections and used to obtain  $T_A$  from (1).

Suppose that the effective sky temperature at right ascension  $\alpha$  and declination  $\delta$  is  $T(\alpha, \delta)$ , and that the normalized power sensitivity of the aerial is  $p(\theta)$ , where  $p(0)=1$  and the angle  $\theta$  is measured from the direction of the aerial. Then

$$T_A \Omega_A = T_A \int_{4\pi} p(\theta) d\omega = \int_{4\pi} T(\alpha, \delta) p(\theta) d\omega, \quad \dots \dots \dots (2)$$

where  $\Omega_A$  is the solid angular beam width of the aerial and  $d\omega$  an element of solid angle. We divide  $\Omega_A$  into two parts  $\Omega_M$  and  $\Omega_S$  where the subscript  $M$  includes the main and first subsidiary lobes of the aerial and extends from  $\theta=0$  to 0.9

radian, and  $S$  includes the smaller lobes from  $\theta=0.9$  to  $\pi$ . We assume the mean maxima of the minor lobes to have  $p=0.014$  and therefore the mean  $p(\theta)$  for the  $S$  part is 0.007. For our aerial we obtain

$$\Omega_M=0.223$$

$$\Omega_S=0.073$$

$$\Omega_A=0.296.$$

The equation (2) then becomes

$$T_A \Omega_A = \int_M T(\alpha, \delta) p(\theta) d\omega + p(\theta) \int_S T(\alpha, \delta) d\omega,$$

where the integrals  $M$  and  $S$  cover the two parts of the aerial reception. For the second term we smooth the minor lobes and regard  $p(\theta)$  as a constant, taking it outside the integral. This term does not vary appreciably as we move the aerial to different parts of the sky, and as we only measure intensity differences it can be neglected. It is convenient to put  $T_A \Omega_A / \Omega_M = T_M$ , where  $T_M$  might be regarded as the effective temperature averaged over the main lobes of the aerial. Then

$$T_M \Omega_M = T_M \int_M p(\theta) d\omega = \int_M T(\alpha, \delta) p(\theta) d\omega. \quad \dots\dots\dots (3)$$

This equation must be solved for  $T(\alpha, \delta)$  from the observed distributions of  $p(\theta)$  and  $T_M$ .

## VI. CORRECTION FOR AERIAL PATTERN

The solution of (3) for  $T(\alpha, \delta)$  gives the correction for the aerial pattern. The usual procedure(3) is to obtain approximate distributions of  $T(\alpha, \delta)$  and apply a numerical integration of (3) to see whether this gives the observed  $T_M$ . The first approximation for  $T(\alpha, \delta)$  is then modified to improve the agreement. It is not possible to obtain details of  $T(\alpha, \delta)$  that are much smaller in angular size than the beam width of the aerial.

The method used here for obtaining the first approximation for  $T(\alpha, \delta)$  was to change  $T_M$  to galactic coordinates  $T_M(l, b)$ . The variation of  $T_M$  with galactic longitude  $l$  is comparatively slow and it is quite evident that most of the aerial correction will be due to the more rapid change with galactic latitude  $b$ . We therefore write (3) in the form

$$T_M(l, b) \int p(\lambda) d\lambda = \int T(l, b-\lambda) p(\lambda) d\lambda, \quad \dots\dots\dots (4)$$

where the variation in galactic latitude alone is taken into account and the variable of integration  $\lambda$  in this direction extends over the main lobes of the aerial. In Section II it was seen that the aerial pattern is very nearly Gaussian with the result that  $p(\theta)=p(\lambda)$ . Now (4) is the familiar equation used for solving the true profile of a spectrum line from the observed and instrumental profile. It can readily be handled by van de Hulst's method(4, 5) if  $T_M$ ,  $T$ , and  $p$  have galactic latitude variations that are Voigt profiles. In the present case it happens that  $p$  is very nearly Gaussian and  $T$  is approximately of dispersion type. These are extreme cases of Voigt profiles. The tables of van de Hulst and Reesinck(5)

may then be applied to correct the profile  $T_M(b)$  and give the true  $T(b)$ . The procedure can then be repeated in the direction of galactic longitude to make the much smaller correction in this direction.

The results so obtained were then put into a numerical integration of (3) to compute  $T_M$  and compare with the observed value. For this purpose the aerial sensitivity pattern was plotted on a movable transparent sheet through which the calculated power distribution plotted in galactic coordinates could be viewed. Allowances were made for the effect of projection and the integration carried only to  $\pm 30^\circ$  galactic latitude. Since there is a high concentration towards the galactic equator, the variation beyond  $b = 30^\circ$  is small, the observed power and the true power converging to the same value.

Trial and error adjustments were then made to improve the agreement, the main adjustment being to increase the peak-value along the galactic equator in the vicinity of the galactic centre.

In Figure 7 we have compared the numerical integration of the finally adopted contours with the observed power values.

## VII. RESULTS

The results were converted into the usual unit watts per sq. metre per (c/s.) per steradian by the black-body equation

$$I = \frac{2kT}{\lambda^2} = 1.23 \times 10^{-23} T.$$

The intensities quoted are therefore total, including both polarizations.

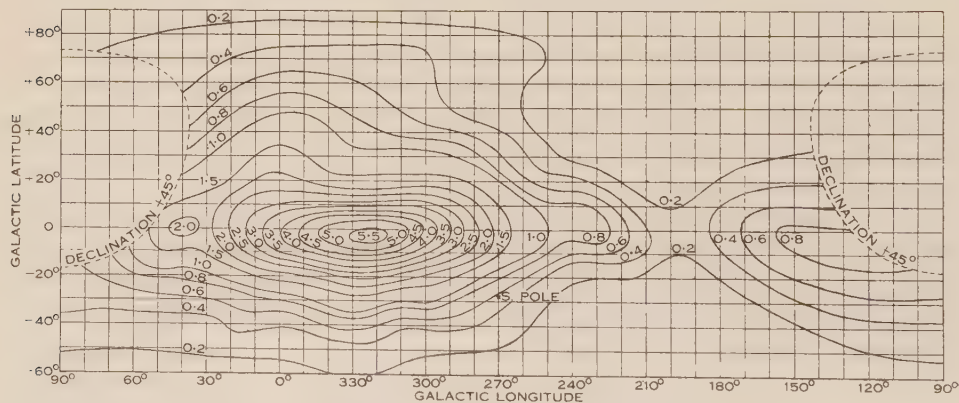


Fig. 5.—Distribution of observed power. Unit =  $10^{-21}$  watts  $\text{m}^{-2}$  (c/s.) $^{-1}$  (steradian) $^{-1}$ .

The measured contours of galactic radio-noise intensity are shown in Figure 5, and the contours corrected for aerial pattern in Figure 6. It has already been mentioned that the radiation from the coldest parts of the sky—probably about  $1.2 \times 10^{-21}$  W.m. $^{-2}$  (c/s.) $^{-1}$  (steradian) $^{-1}$ —must be added to these values.

There may be unresolved detail near the maximum at galactic longitude  $l = 322^\circ$ , latitude  $b = -3^\circ$ , and at this point errors of the order of 50 per cent. are possible. For most of the diagram the error in intensity should be in the region

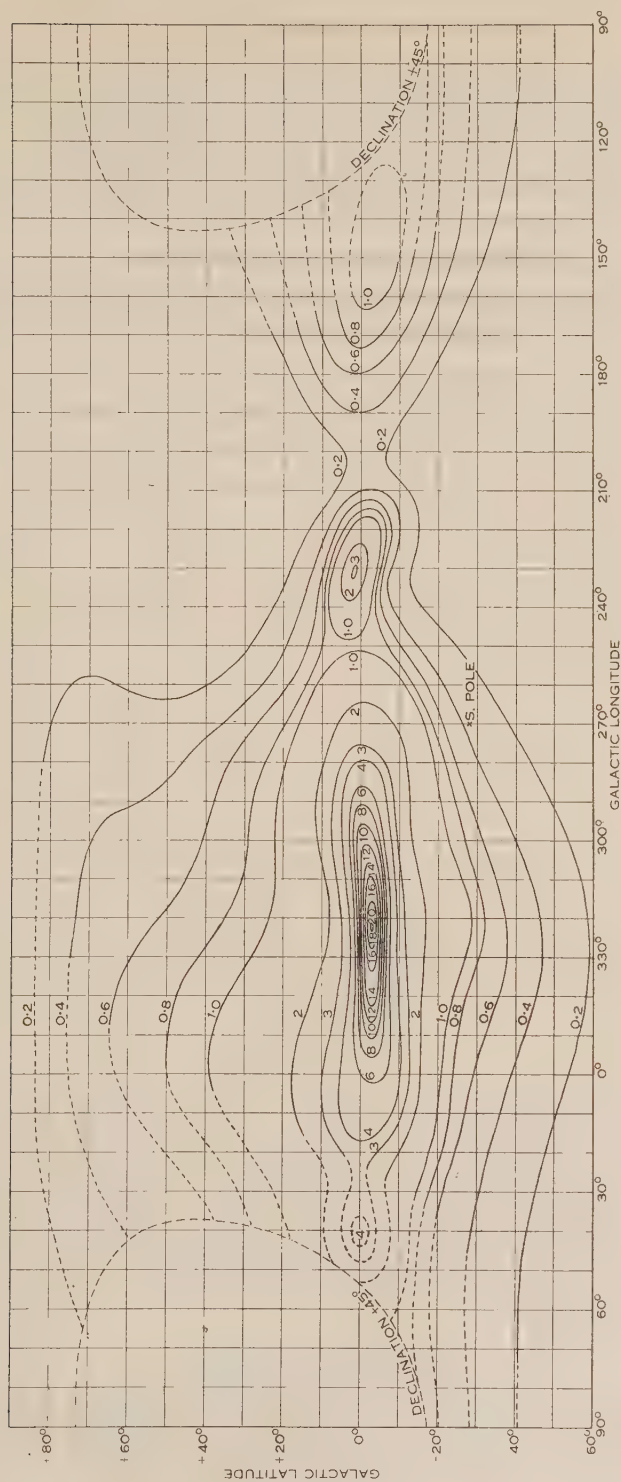


Fig. 6.—Distribution of galactic electromagnetic radiation at 200 Mc/s. (Minimum of sky taken as zero.) Unit =  $10^{-21}$  watts m. $^{-2}$  (c/s.) $^{-1}$  (steradian) $^{-1}$ .



of 20 per cent. The probable error of location is about  $1^\circ$ . The correction for aerial pattern becomes indeterminate close to the horizon. All values north of declination  $+25^\circ$ , which are shown as broken lines in Figure 6, are extrapolated and follow the general trend of the observed contours.

The intensities of the discrete sources have not been subtracted from the data. Most of them, however, are too small to influence the contours. No radiation from the Magellanic clouds was detectable.

For standardizing solar-noise measurements it is convenient to use galactic noise intensities in right ascension and declination coordinates, and not corrected for aerial pattern. Such intensities are shown in Figure 4, where the unit of

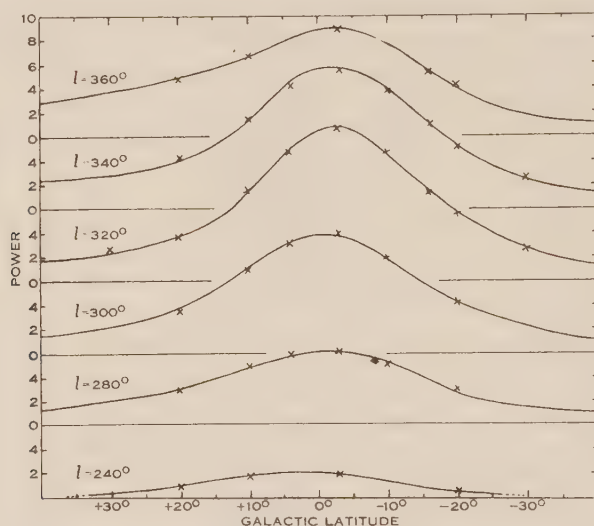


Fig. 7.—Comparison of observed power with that calculated from the distribution shown in Figure 6. Unit =  $10^{-22}$  watts  $\text{m}^{-2}$   $(\text{c/s.})^{-1}$ .

— Observed power.  
 × Calculated power.

measured intensity is  $I_M \Omega_M$ . These values are directly comparable with solar intensities in the usual watts  $\text{m}^{-2}$   $(\text{c/s.})^{-1}$ . Accurate measurements of the galactic intensity in this form could provide a permanent and convenient noise standard for comparing solar observations with the present aerial. The results in Figure 4 might be regarded as a preliminary estimate of this standard. Some modification would be required if the values were to be used with aerials having different beam widths.

### VIII. ACKNOWLEDGMENTS

This work has been carried out at the Commonwealth Observatory, Canberra, A.C.T. The authors have had the benefit of discussion with Dr. J. L. Pawsey and Mr. J. G. Bolton, Division of Radiophysics, C.S.I.R.O.

## IX. REFERENCES

- (1) ALLEN, C. W.—*Mon. Not. R. Astr. Soc.* **107** : 386 (1947).
- (2) PAWSEY, J. L., and MCCREADY, L. L.—The measurement of solar radio-frequency radiation. C.S.I.R.O. Aust. Radiophysics Rep. No. 74 (1948).
- (3) HEY, J. S., PARSONS, S. J., and PHILLIPS, J. W.—*Proc. Roy. Soc. A* **192** : 425 (1948).
- (4) VAN DE HULST, H. C.—*Bull. Astr. Inst. Neth.* **9** : 225 (1941) ; **10** : 79 (1946).
- (5) VAN DE HULST, H. C., and REESINCK, J. J. M.—*Astrophys. J.* **106** : 121 (1947).

# GALACTIC RADIATION AT RADIO FREQUENCIES

## II. THE DISCRETE SOURCES

By G. J. STANLEY\* and O. B. SLEE\*

[*Manuscript received November 4, 1949*]

### *Summary*

An account is given of observations on a number of discrete sources of radio-frequency radiation, together with a brief description of the observational techniques used in their detection, and of the methods employed to determine their positions and angular extent.

Noise spectra of four of the more intense sources have been measured over a frequency range of 40–160 Mc/s.: in three cases the change of intensity with wavelength is found to be greater than that of the background continuum, and in one case less. Two of the sources have been provisionally identified with astronomically rare objects, the Crab Nebula and N.G.C. 5128.

Finally, the short period fluctuations in the intensity of some of the sources, notably that in Cygnus, are discussed. Evidence is presented which, contrary to previous views, suggests that these fluctuations are of terrestrial rather than of extra-terrestrial origin.

## I. INTRODUCTION

In 1946, whilst making a survey of galactic noise at 64 Mc/s., Hey, Parsons, and Phillips(1) noticed that the noise intensity from the constellation of Cygnus was by no means constant. They observed fluctuations of about 15 per cent. of the total power received by their aerial system, the average period of the fluctuations being about one minute. By direction finding they localized the source of these fluctuations to an area of  $2^\circ$  centred at Right Ascension 20 hours, Declination  $+43^\circ$ , and suspected the existence of other active areas within  $8^\circ$  of this position. Bolton and Stanley(2), using an interference technique with the sea as a reflector, showed that the fluctuations originate in a discrete source less than 6 minutes in angular width, and later reported the existence of further discrete sources(3), apparently constant in intensity. Both teams suggested a possible similarity between the Cygnus fluctuations and enhanced solar radiation associated with active sunspots. Ryle and Smith(4), however, using a vertical interferometer, showed that the radiation from the Cygnus source was unpolarized, in contrast with that from active sunspots. They, too, found a number of other discrete sources.

The first part of this paper discusses the detection of extremely small signals, methods of observing discrete galactic sources, and the determination of their positions and angular widths. The second part gives details of observations on discrete sources made at this Laboratory, including the variation of

\* Division of Radiophysics, C.S.I.R.O., University Grounds, Sydney.

intensity with frequency (over the range 40–160 Mc/s.) for several of the larger sources. Finally, short-period fluctuations in the intensity of the sources are discussed, with particular reference to two years' observations on the Cygnus source, including spaced aerial measurements between Australia and New Zealand.

## II. OBSERVATIONAL TECHNIQUES

### (a) *Receiver Sensitivity*

Over the range of frequencies considered, galactic noise received on a moderately directive aerial system is of the same order as the noise generated in the receiver. The radiation from individual discrete sources is only a small fraction of the incident noise. The limit of sensitivity of a receiver capable of detecting signals from these small sources is set by the fluctuation level of the total noise observed at the detector. Contributions to the fluctuation level arise from :

- (i) Statistical fluctuation of the noise generated in the first stages of the receiver.
- (ii) Fluctuations of receiver gain due to variation in filament and high-tension supplies.
- (iii) Terrestrial noise.

(i) *Statistical Fluctuations of Noise in the Receiver*.—The theoretical limit of receiver stability is determined by the statistical fluctuations of noise generated in the first stages of the receiver. Dicke(5) has shown that, in terms of the aerial temperature,\* this limit is given by the formula

$$\frac{\Delta T}{T} = \frac{\pi^{3/2} N}{8} \left( \frac{\alpha}{\Delta \omega} \right)^{\frac{1}{2}},$$

where  $N$  is the noise factor of the receiver,

$\Delta \omega$  is the bandwidth of the receiver (radians per second),

$\alpha$  is the reciprocal of the time constant used in the recording system,

$T$  is the aerial temperature,

$\Delta T$  is the minimum observable temperature change.

For a typical 100 Mc/s. system, where  $N=3$ ,  $\alpha=0.2$ , and  $\Delta \omega=2\pi \times 10^6$ ,

$$\Delta T = 0.1^\circ \text{K.}$$

for an aerial at ambient temperature.

(ii) *Supply Voltage Variations*.—It has been found that, with careful attention to the mechanical and electrical layout of the receiver, the only significant changes in receiver output from internal causes are due to variations in gain resulting from fluctuations in the high-tension and filament voltages. For an average 100 Mc/s. receiver a change of 1 part in 3000 in high-tension voltage or 1 part in 1500 in filament voltage produces a change of 1 part in 1000 in second detector voltage. The optimum stability attained in the present

\* The concepts *aerial temperature* and *equivalent black-body temperature* have been described in detail in Part I of this series.



observations was 1 part in 10,000 in high-tension and 1 part in 6000 in filament supplies. For an aerial at ambient temperature this represented a minimum observable temperature change of about  $0.3^\circ\text{K}$ . In the receivers used, the second detector was followed by a direct-current amplifier in which most of the detector output, representing galactic noise plus receiver noise, was balanced out and the remainder presented on a recording milliammeter.

In an alternative system used by Dicke for observation of microwave radiation, the receiver input is switched alternately between the aerial and a matched load at ambient temperature. As the signal is only received for half the time and then integrated, the theoretical limit of sensitivity, as given by the formula quoted above, will be  $\sqrt{2}\Delta T$ . The practical limits of such a system are unknown. To improve appreciably on the sensitivity of the Dicke system, supply voltages for the previously described system would have to be regulated to the order of 1 part in  $10^5$ .

(iii) *Terrestrial Noise Sources*.—Radio-frequency noise, due to distant atmospherics, car ignition, transmission line corona, etc., sets a practical limit to the sensitivity of any system. In the present investigation the fluctuation level due to terrestrial noise, even during the optimum observing period between 2 a.m. and dawn, was generally greater than that arising within the equipment. Sensitivity close to the practical limit discussed earlier has been achieved in field observations from selected sites.

### (b) Interference Technique

Owing to the difficulty of constructing highly directional aerial systems, discrete sources of radio-frequency energy have only been recognized against the general noise background by using some form of interference technique. Two

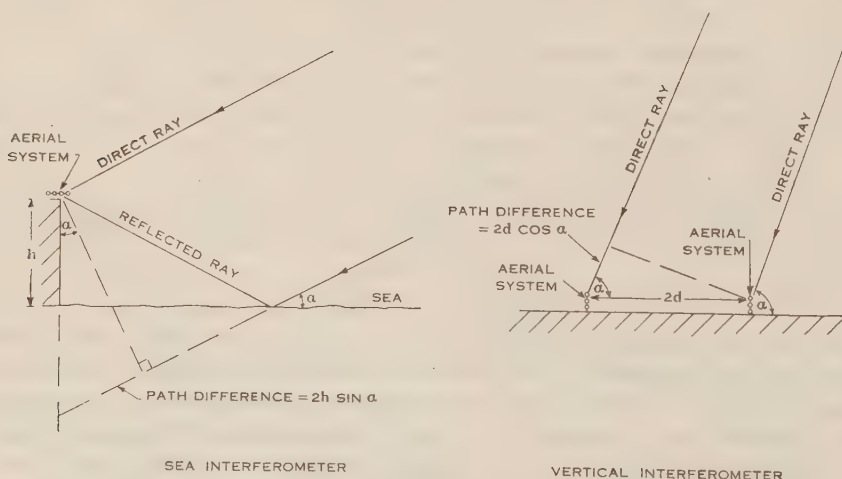


Fig. 1.—Schematic diagram of sea and vertical interferometer systems.

forms of interferometer are in use at present. One, an adaptation of "Lloyd's mirror" using the sea as a reflector, was first used by McCready, Pawsey, and Payne-Scott(6) for studying solar radiation at 200 Mc/s., and later by Bolton and

Stanley(2) for the detection of discrete sources. The second, similar to the Michelson stellar interferometer and consisting of two spaced aerials, was first used by Ryle and Vonberg(7) for observation of solar radiation at 80 and 175 Mc/s. The principles of the two interferometers are illustrated in Figure 1.

The advantages of the sea interferometer used in the present investigations are :

- (i) The same sensitivity is achieved with a single aerial unit.
- (ii) The source first appears sharply in contrast to the general background at rising, whereas with the vertical system a slowly increasing amplitude pattern is obtained.
- (iii) The sharp appearance at rising is useful in discriminating between a number of sources close together.

#### (c) *Measurement of Angular Width*

It has been shown elsewhere(6, 8) that an estimate of the angular width of the source in a direction at right angles to the horizon can be made from an examination of sea interferometer patterns. If  $R$  is the ratio of the received power (above the extrapolated background level) at the maxima and minima of the interference pattern,  $h$  the height of the aerial above sea-level, and  $\lambda$  the wavelength used,  $W$  the effective angular width is given by

$$W = \frac{\lambda}{\pi h} \sqrt{3R}.$$

In practice, an upper limit is generally set by the noise fluctuation level of the receiver, which affects the measurement of the heights of the minima. It may appear that unlimited resolution can be obtained by merely increasing the height of the observing site, but this is limited by two factors. One is the "effective sea reflection coefficient", which is less than unity for altitudes below about  $1^\circ$ , due to the earth's curvature. The second is the bandwidth of the receiver, which increases the height of the minima and decreases the height of the maxima with increasing fringe order. Although both these effects can be computed (see Appendices I and II), some uncertainty must remain.

It is obvious that the resolving power of the equipment depends principally on the intensity of the source, and thus accurate measurements have only been made for the more intense sources.

#### (d) *Determination of Position*

For the determination of position in the case of the sea interferometer, each minimum of the patterns obtained at rising or setting represents the time at which the source attains a certain altitude (determined by instrumental and site constants and a knowledge of atmospheric refraction).

The intersection on the celestial sphere of two lines, one of which is determined by the sidereal time of rising and the other by the declination, gives the position of the source. The line of declination is estimated from the periodicity of the interference pattern or by direction finding when an insufficient number of minima is available.

TABLE 1  
LIST OF CONFIRMED SOURCES

Number and Year of Discovery	Temporary Designation	Intensity at 100 Mc/s. (watts m <sup>-2</sup> (c/s.) <sup>-1</sup> ) 10 <sup>-24</sup>	Upper Limit of Angular Width (minutes)	Method of Position Finding	Celestial Coordinates (epoch 1948)	Galactic Coordinates	Number of Observations	Remarks
1—1946	Cygnus-A	125	1.5	R and S	R.A. 19 hr. 58 min. 14 sec. Dec. +40° 38'	<i>l</i> 44° <i>b</i> +4½°	320	Amplitude fluctuation first observed by Hey, Parsons, and Phillips(1)
2—1947	Taurus-A	18.5	6	R and S	R.A. 05 hr. 31 min. 30 sec. Dec. +22° 01'	<i>l</i> 150° <i>b</i> -4°	80	Possibly M 1 (N.G.C. 1952), the Crab Nebula
4—1947	Virgo-A	12.5	5	R and S	R.A. 12 hr. 28 min. 06 sec. Dec. +12° 41'	<i>l</i> 260° <i>b</i> +74°	100	Possibly M 87 (N.G.C. 4486)
5—1947	(Centaurus)	(8)	30				10	See Subsection III (f)
6—1947	Centaurus-A	18.5	7	R and S	R.A. 13 hr. 22 min. 20 sec. Dec. -42° 37'	<i>l</i> 274° <i>b</i> +20°	50	Possibly N.G.C. 5128
7—1948	Hercules-A	2	30	R, L of R, D	R.A. 16 hr. 50 min. Dec. +5°	<i>l</i> 351° <i>b</i> +26°	30	See position map I, Fig. 2
8—1948	Taurus-C	3	15	L of R, D	R.A. 04 hr. 38 min. Dec. +28°	<i>l</i> 140° <i>b</i> -10°	15	See position map II, Fig. 2, also Reference (8) for detailed information
9—1948	Taurus-B	6	30	L of R	R.A. 05 hr. 32 min. Dec. +24°	<i>l</i> 151° <i>b</i> -8°	15	See position map II, Fig. 2, also Reference (8) for detailed information
10—1948	Fornax-A	2	15	R, L of R, D	R.A. 03 hr. 11 min. Dec. -36°	<i>l</i> 204° <i>b</i> -57°	15	See position map III, Fig. 2
11—1948	Serpens-Cauda-A	3	15	R, L of R, D	R.A. 18 hr. 43 min. Dec. -75°	<i>l</i> 5° <i>b</i> +11°	6	See position map IV, Fig. 2
12—1948	(Centaurus)	2	30			<i>l</i> 265½° <i>b</i> +17°	4	Rises before Centaurus-A, almost circumpolar at the latitude of the observing site
15—1948	(Leo)	1	30	L of R	R.A. 11 hr. 52 min. Dec. +17°	<i>l</i> 223° <i>b</i> +74°	5	Rises just before Virgo-A. See map V, Fig. 2
17—1949	Scorpius-A	2	30			<i>l</i> 320° <i>b</i> -10°	5	See position map VI, Fig. 2
18—1949	Serpens-Cauda-B	2	30	L of R, D	R.A. 18 hr. 11 min. Dec. 15°	<i>l</i> 343° <i>b</i> -1°	5	See position map VII, Fig. 2
19—1949	Sextans-A	2	30	L of R, D	R.A. 09 hr. 55 min. Dec. -5°	<i>l</i> 214° <i>b</i> +38°	4	See position map VIII, Fig. 2. On the border of Comaba and Caelum
20—1949		2	30	L of R, D	R.A. 05 hr. 01 min. Dec. -36°	<i>l</i> 205° <i>b</i> -85°	10	See position map IX, Fig. 2
21—1949	Puppis-A	3	30	R, L of R, D	R.A. 08 hr. 18 min. Dec. -42°	<i>l</i> 227° <i>b</i> -2°	8	See position map X, Fig. 2
22—1949	Pictor-A	3	30	R, L of R, D	R.A. 05 hr. 18 min. Dec. -44°	<i>l</i> 217° <i>b</i> -33°	6	See position map X, Fig. 2

In practice, however, this method is somewhat inaccurate owing to the low reading-accuracy of the times of occurrence of minima and the considerable scatter between results on different occasions (presumably due to refraction). A higher order of accuracy can be obtained if the source is observed at both rising and setting from the same or different sites. This method was used by Bolton and Stanley in determining the position of the source Taurus-A, and is described in detail elsewhere(8).

### III. THE DISCRETE SOURCES

#### (a) *General*

Details of the known discrete sources are set out in Table 1. The larger sources, discovered late in 1947 and early in 1948 with the use of only small aerial systems, have been studied extensively. The smaller sources have been found by use of the large array described in Part I of this series. Owing to the relatively small beam width of this system, systematic surveying is necessarily a lengthy process. As the full sensitivity of the system could only be realized for a few hours before dawn, small sources detected during the optimum observing period could not always be confirmed by later observations. Sources of which less than three records are available have therefore not been included in the table.

Accurate position finding observations have only been made for several of the larger sources. Approximate positions (shown in brackets) have been found by a number of methods designated by the following letters in Table 1 :

*R*, from observations of the track of the source above the horizon after rising.

*R* and *S*, from observations of the source at rising and setting.

*L* of *R*, from the time of rising which determines a line on the celestial sphere.

*D*, by direction finding at rising.

With the exception of the major sources in Taurus, Virgo, and Centaurus, no great accuracy is claimed for these measurements; probable limits in the positions of the sources are shown on the series of position maps (Fig. 2).

The angular widths given refer to measurements at 100 Mc/s., as the sensitivity of the equipment was greater at this frequency. Upper limits of 30 minutes of arc are given where no accurate measurements have been made; it is considered unlikely that such sources would have been seen with the interferometer if their angular widths had exceeded this figure.

Estimates of intensity similarly refer to 100 Mc/s. The accuracy of these measurements varies from  $\pm 10$  per cent. for the larger sources to  $\pm 50$  per cent. for the very weak sources.

#### (b) *Cygnus-A*

The position of this source, given in Table 1, is the result of a long series of measurements made from the north-east coast of New Zealand in June and July 1948. Almost the complete track of the source above the northern horizon was observed—the reflected ray was intercepted only for a short period between



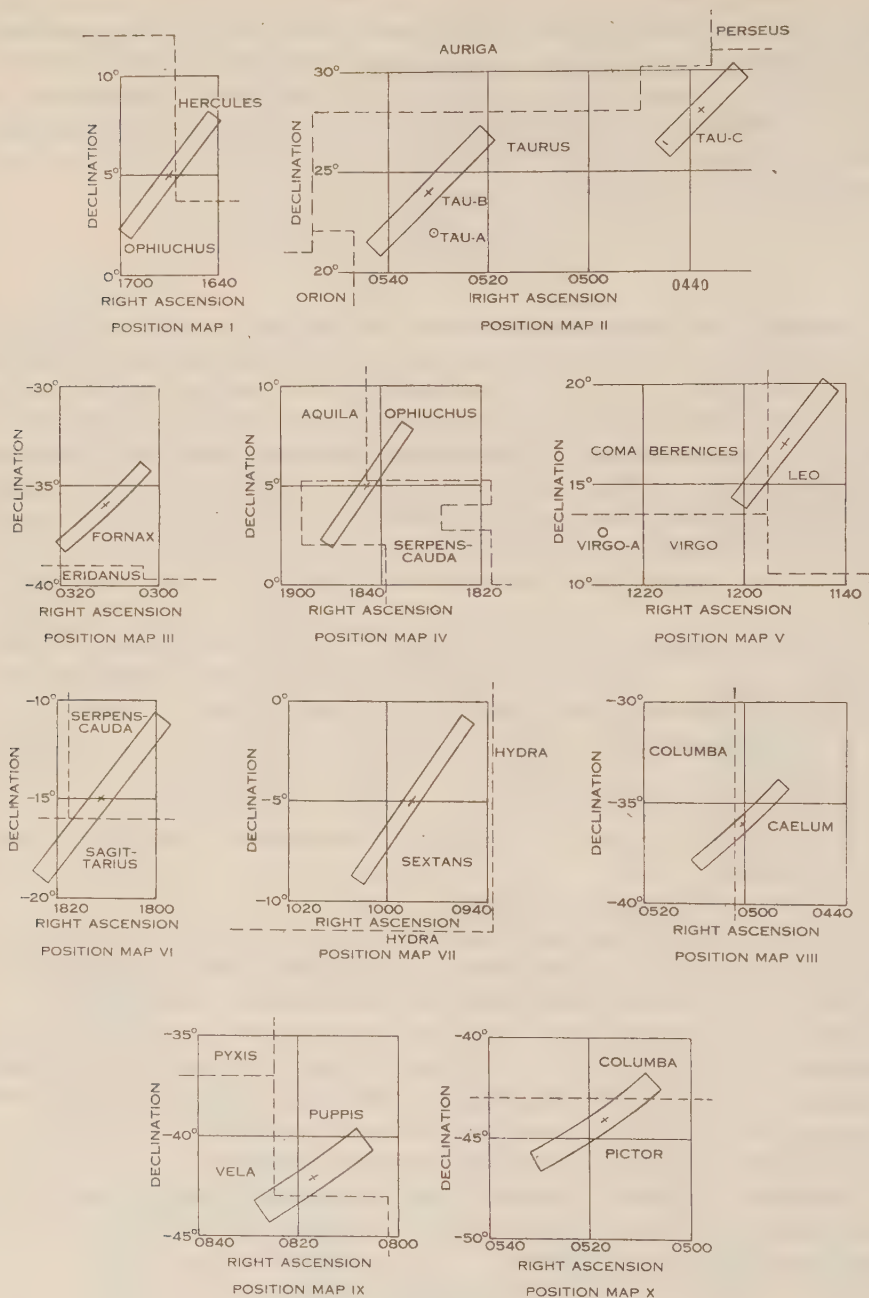


Fig. 2.—Position maps for a number of discrete sources. Full lines mark limits of position and dotted lines boundaries of constellation.

culmination and setting of the source by a group of islands. The Right Ascension was found from comparison of the time of corresponding minima on each side of culmination and the declination from the semi-diurnal arc and the altitude of the source at culmination. The main feature of these observations was the great inconsistency between the deduced position on different nights and between observations at different periods of the same night. The authors have been unable to reconcile their results with those of Ryle (personal communication) and also of Mills of this Laboratory using vertical interferometers. It is unlikely, however, that the source lies outside the area bounded by Right Ascension 19 hr. 58 min.–20 hr. 00 min., and Declination  $+40^{\circ} 20' - 40^{\circ} 40'$ .

(c) *Taurus-A*

The position of this source closely coincides with that of the Crab Nebula (Plate 1), the expanding shell of an old supernova explosion. This object appears to be a likely source of high level radio-frequency noise. For a source size of five minutes the intensity at 100 Mc/s. corresponds to an equivalent temperature of about 2,500,000 °K. The present estimates of density and temperature in the Crab Nebula fall well short of explaining this result by strictly thermal processes. Non-thermal components resulting from turbulence within the nebula do not, however, seem unlikely.

(d) *Centaurus-A*

This source has been provisionally identified with the extra-galactic nebula N.G.C. 5128 (Plate 1). The fact that features of this nebula—the gas emission spectrum, the high density of interstellar matter, and the anomalous position of the obscuring band—are already known to be unusual suggests the possibility that its radio-frequency emission may also be unusual, and hence lends some support to the identification. (Recent work by Evans(9) raises the possibility that N.G.C. 5128 is not extra-galactic. He considers that this object is a gas cloud illuminated by a point at its centre. The dark marking across the nebula is presumably due to obscuring matter in the line of sight.)

(e) *Temporary Phenomena*

One example of a possible temporary source is 5·47 (probably in Centaurus, see Table 1). Some 10 records of this source at rising were obtained in November and December 1947. No further records were taken until June 1948, when no trace of the source could be found. Subsequent attempts to obtain records with higher sensitivity have also failed. There have been several other cases when small sources have been detected during systematic surveys—sometimes on several nights in succession—and attempts to secure additional confirmation at a later date have not succeeded. The possibility of receiver fluctuations accounting for such phenomena can be almost entirely excluded as the interference patterns obtained had lobe periods consistent with the bearing of the aerial. A much longer period of observation, however, will be necessary before the existence of temporary sources can be established beyond doubt.

*(f) Distribution of Sources*

The known sources are shown in Figure 3 on an equal-areas chart in galactic coordinates. The two dotted lines enclose the areas that are almost circumpolar to the present site, and thus have not been surveyed. The size of the dots gives a rough idea of the relative intensity of the sources. The number of sources available at present is too small to make any deductions from this distribution, but there seems a general tendency to crowd towards the galactic equator, and perhaps a noteworthy absence of sources between longitudes 330 and 210° south of the equator.

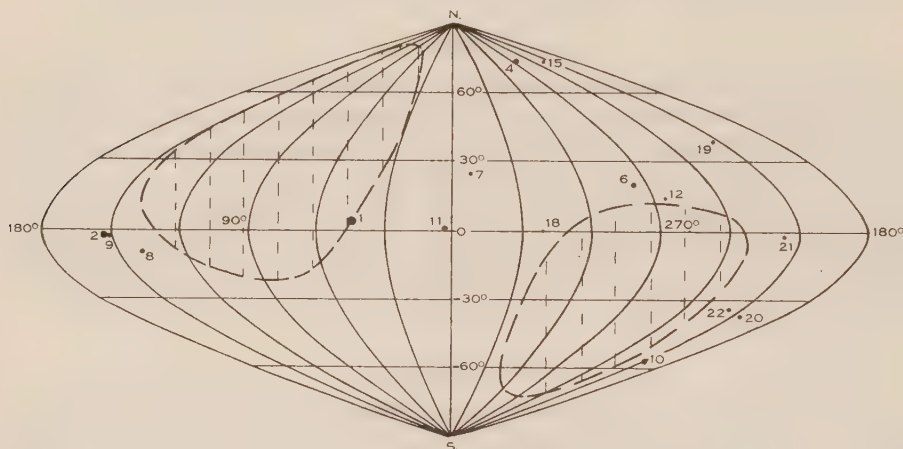


Fig. 3.—Distribution of discrete sources on an equal areas galactic chart. Numbers refer to Table I.

## IV. NOISE SPECTRA

In Figures 4 (a) to 4 (d) noise "spectra" of the four most intense sources that can be observed from the present site are shown. Measurements were made on Cygnus and Taurus-A at five frequencies, 40, 60, 85, 100, and 160 Mc/s. The measurements on the sources Virgo-A and Centaurus-A were made during daylight when the terrestrial noise fluctuations were too great to obtain results at 40 Mc/s. A harmonic of a C.W. transmission prevented measurements of the intensity of Centaurus-A at 85 Mc/s.

Although the accuracy of these observations was not high, it was sufficient to show differences in the spectra of the various sources. For three of the sources the intensity is approximately inversely proportional to frequency, whereas for Taurus-A the intensity is apparently independent of frequency over the range of frequencies considered. The latter would be consistent with "thermal" radiation from an optically thin shell.

Figure 5 shows spectra of the north galactic pole and of a region round the galactic centre obtained with a single Yagi aerial. These are given in terms of equivalent temperature and the additional points at 40, 64, and 90 Mc/s. are due to Moxon(10) and Hey, Parsons, and Phillips(1). For the north galactic

pole the equivalent temperature varies as  $f^{-2.35}$  and for the galactic centre as  $f^{-2.65}$  (in terms of intensity  $f^{-0.35}$  and  $f^{-0.65}$ ). It is interesting to note that, if a possible variation of source size with frequency can be ignored, for three of the discrete sources the intensity would increase more rapidly with decreasing frequency than that of the integrated background noise.

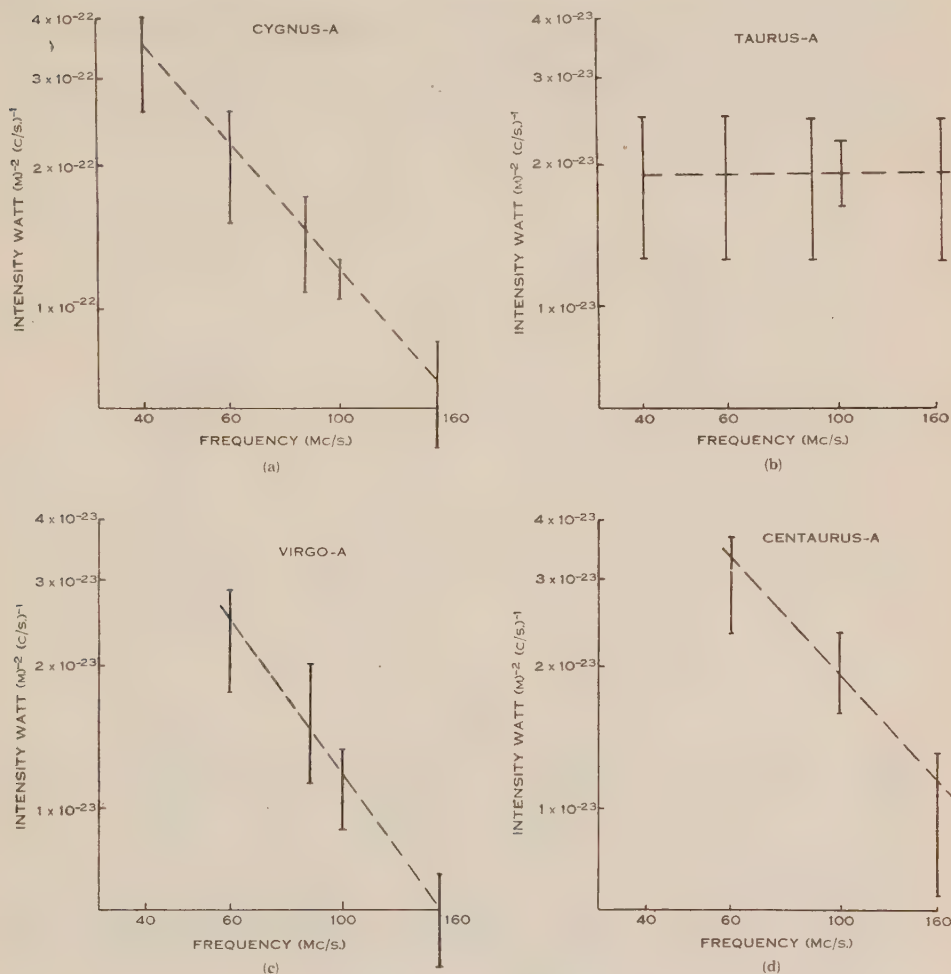


Fig. 4. (a) to (d)—Noise "spectra" of four discrete sources. Vertical lines indicate the range of observed values of intensity.

## V. FLUCTUATIONS

Since their discovery by Hey, Parsons, and Phillips(1) the fluctuations in the received intensity of the source in Cygnus have remained one of the puzzling features of this phenomenon. It was at first thought that the radiation from the source consisted of two distinct components(2) as in the case of thermal and enhanced radiation from the sun. However, a longer period of observations



has shown that this view is not correct, and that the mean intensity of the source is constant to within 10 per cent. Present knowledge of the fluctuations may be summarized as follows:

- (i) The probability of fluctuations in intensity decreases with increasing radio frequency. Fluctuations have only rarely been observed at 160 Mc/s.
- (ii) The amplitude (relative to the mean level) of the fluctuations increases with decreasing radio frequency.
- (iii) The average period of the fluctuations decreases with decreasing radio frequency.

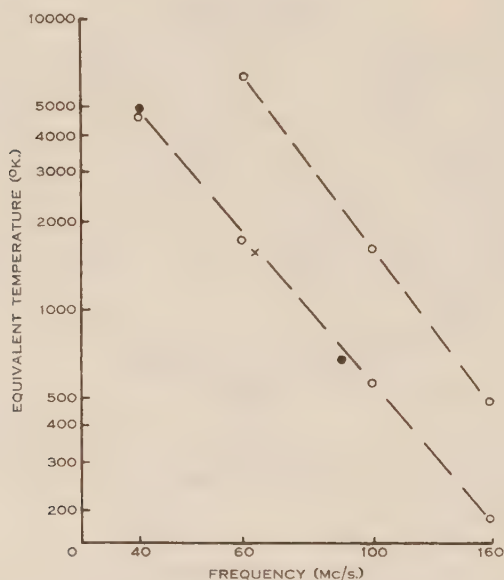


Fig. 5.—Variation of aerial temperature with frequency for the north galactic pole (lower curve) and a region surrounding the galactic centre (upper curve).

○ Radiophysics observations; ● Moxon; × Hey.

- (iv) There is general correspondence between the degree of fluctuation on a number of radio frequencies observed simultaneously from the same site.
- (v) There is no detailed correspondence between fluctuations on frequencies as close as 85 and 100 Mc/s. observed simultaneously from the same site.
- (vi) There is some evidence of an annual variation in the occurrence of the fluctuations. The data in Table 2 refer to observations on the source of approximately 1 hour's duration immediately after its rising. From each record the amplitude of the fluctuations is assessed as a percentage of the mean intensity. The monthly means of the values obtained are plotted in Figure 6. The maximum degree of fluctuation occurs

between July and September when the source rises in the early evening, and the minimum in February and March when the source rises about sunrise. Observations, however, commencing before and finishing after sunrise showed that no appreciable change in the character of the fluctuations occurred after the sun rose.

TABLE 2  
FLUCTUATIONS OF SOURCE IN CYGNUS

Month	Mean Monthly Average of Amplitude of Fluctuations as Per Cent. of Mean Intensity	Number of Observations
1947		
July .. ..	60	10
December .. ..	55	10
1948		
January .. ..	42	8
February .. ..	—	—
March .. ..	17	12
April .. ..	32	21
May .. ..	51	23
June .. ..	56	23
July .. ..	61	23
August .. ..	49	11
September .. ..	41	26
October .. ..	52	18
November .. ..	58	22
December .. ..	58	10
1949		
January .. ..	62	5
February .. ..	20	3
March .. ..	11	16
April .. ..	32	7
May .. ..	51	15
June .. ..	44	17

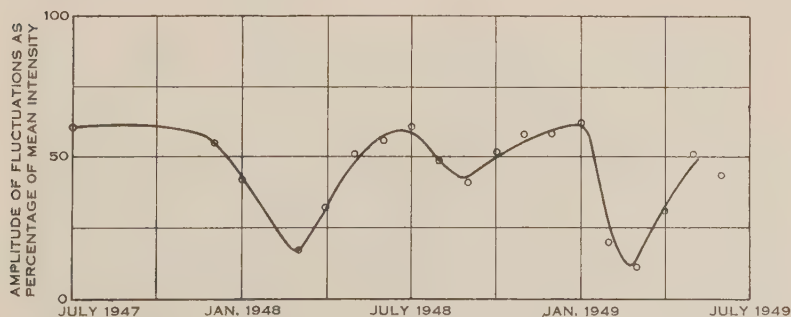


Fig. 6.—Degree of fluctuation of the intensity of the source in Cygnus.  
Individual points are monthly means.

Records of the source taken simultaneously from sites on the east coasts of Australia and New Zealand showed no correlation in the fluctuations, whereas similar observations of enhanced solar radiation showed a 90 per cent. correlation. This indicates that the fluctuations are either extraordinarily directive or "locally" impressed.

Fluctuations in intensity have been observed for a number of the smaller sources, particularly Centaurus-A. The degree and incidence of fluctuation, however, is considerably less than that of the Cygnus source but is believed to follow the same annual curve. Fluctuations are most noticeable in all cases when the altitude of the source is less than  $2^\circ$ . It should be pointed out that observation of the very small sources depends on some degree of integration (Subsection II (a)) which would tend to obscure fluctuations.

There are four features of the fluctuation phenomenon which would suggest an origin in the ionosphere:

- (i) The spaced aerial observations on Cygnus.
- (ii) The annual variation in the Cygnus fluctuations.
- (iii) The increasing degree of fluctuation with decreasing frequency.
- (iv) Marked fluctuations in most cases at low altitude.

The absence of fluctuations in some of the minor sources may be due to angular width. If the fluctuations are due to irregularities in the ionosphere it is likely that the effect would be more pronounced in the case of sources of small angular width. Recent measurements indicate that the source in Cygnus has an angular width of less than  $1' 30''$  whereas the angular widths of a number of the minor sources are believed to be several minutes of arc. This suggestion may be compared with the "twinkling" of stars and the steady appearance of planets.

## VI. REFERENCES

- (1) HEY, J. S., PARSONS, S. J., and PHILLIPS, J. W.—Fluctuations in cosmic-radiation at radio frequencies. *Nature* **158**: 234 (1946).
- (2) BOLTON, J. G., and STANLEY, G. J.—The variable source of radio-frequency radiation in the constellation of Cygnus. *Nature* **161**: 312 (1947).
- (3) BOLTON, J. G.—Discrete sources of galactic radio-frequency noise. *Nature* **162**: 141 (1948).
- (4) RYLE, M., and SMITH, F. G.—A new intense source of radio-frequency radiation in the constellation of Cassiopeia. *Nature* **162**: 462 (1948).
- (5) DICKE, R. H.—The measurement of thermal radiation at microwave frequencies. *Rev. Sci. Instrum.* **1946**: 268 (1946).
- (6) MCCREADY, L. L., PAWSEY, J. L., and PAYNE-SCOTT, RUBY.—Solar radiation at radio frequencies and its relation to sunspots. *Proc. Roy. Soc. A* **190**: 357 (1947).
- (7) RYLE, M., and VONBERG, D. D.—An investigation of radio-frequency radiation from the sun. *Proc. Roy. Soc. A* **193**: 98 (1948).
- (8) BOLTON, J. G., and STANLEY, G. J.—The position and probable identification of the galactic source Taurus-A. *Aust. J. Sci. Res. A* **2**: 139 (1949).
- (9) EVANS, D. S.—Photometry of N.G.C. 5128. *Mon. Not. R. Astr. Soc.* **109**: 94 (1949).
- (10) MOXON, L. A.—Variation of cosmic radiation with frequency. *Nature* **158**: 758 (1946).

## APPENDIX I

*The Effect of Receiver Bandwidth on Interference Patterns*

For the case of the sea interferometer described in Subsection II (a) the phase difference between the direct and reflected rays is  $D + \pi$ , where

$$D = \frac{4\pi h \sin \alpha}{\lambda} = \frac{4\pi h f \sin \alpha}{c} \dots\dots\dots (1)$$

If  $P_o$  is the mean power in the band  $\Delta f$  intercepted from the direct beam, the mean power intercepted from the two interfering beams in the elementary band  $df$  is

$$2P_o(1 - \cos D) \frac{df}{\Delta f}.$$

The mean power over the band  $\Delta f$ , obtained by integrating between the limits  $f \pm \frac{\Delta f}{2}$ , is then

$$P = 2P_o \left\{ 1 - \cos D \frac{\sin \left( \frac{D \Delta f}{2f} \right)}{D \frac{\Delta f}{2f}} \right\}, \dots\dots\dots (2)$$

which reduces to the elementary case

$$P = 2P_o(1 - \cos D), \dots\dots\dots (3)$$

if  $\Delta f$  is sufficiently small.

Under actual conditions (e.g.  $\frac{\Delta f}{f} = \frac{1}{100}$ ) the stationary values of  $P$  given by (2) are not sensibly different from those of (3). Thus minima occur where

$$D = 2r\pi,$$

and maxima where

$$D = (2r + 1)\pi \quad r = 0, 1, \dots\dots\dots (4)$$

Thus for the  $r$ th lobe

$$\frac{P_{min}}{2P_o} = 1 + \frac{\sin \left( (r + \frac{1}{2})\pi \Delta f / f \right)}{(r + \frac{1}{2})\pi \Delta f / f}$$

and

$$\frac{P_{max}}{2P_o} = 1 - \frac{\sin (r\pi \Delta f / f)}{r\pi \Delta f / f} \dots\dots\dots (5)$$

For the early lobes these values approximate to those of the elementary case, but as  $r$  increases they both tend to unity. Figure 7 shows the variation of

$\frac{P_{max} - P_{min}}{4P_o}$  for values  $\lambda = 3$  m.,  $h = 300$  m., and  $\Delta f = 1$  Mc/s.



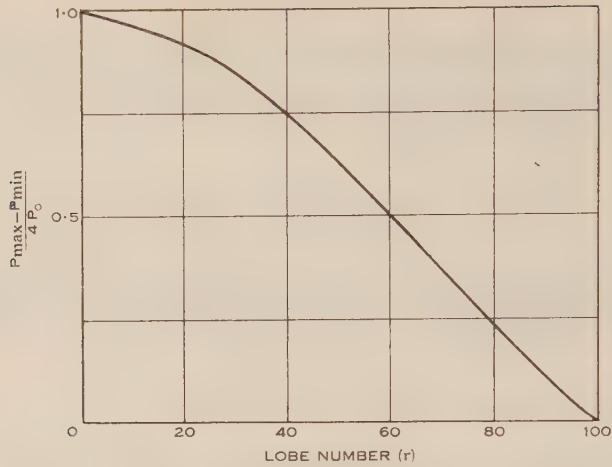


Fig. 7.—Illustration of the effect of receiver bandwidth on interference maxima and minima for numerical examples in Appendix I.

## APPENDIX II

### *The Effect of Earth's Curvature on Sea Interferometer Patterns*

The effect of the curvature of the earth is to reduce the power received by the aerial from the reflected beam (see Fig. 8). This results in incomplete interference.

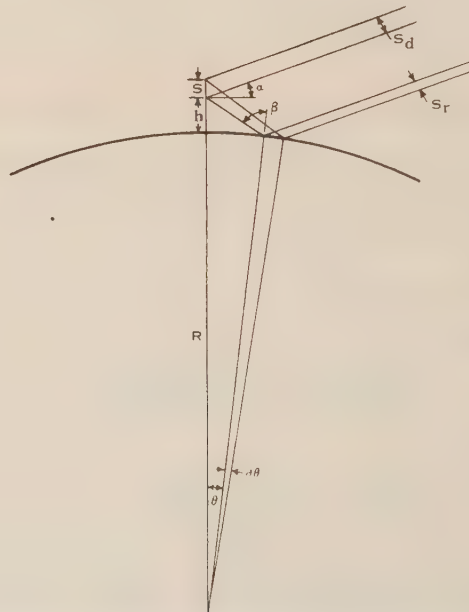


Fig. 8.—Effect of the earth's curvature on the reflected beam.



Fig. 1.—The nebula N.G.C. 5128.



Fig. 2.—The Crab Nebula.



If  $s$  is the "effective height" of the aerial, the widths of the direct and reflected beams intercepted from a source at altitude  $\alpha$  are

$$s_\alpha = s \cos \alpha,$$

and

$$s_n = R\partial\theta \sin(\alpha + \theta), \dots\dots\dots (1)$$

where  $\partial\theta$  is the range of  $\theta$  corresponding to the height  $s$ .

Since the effect occurs only at low altitudes the relation(2)

$$\alpha = \frac{h}{R\theta} - \frac{3\theta}{2} \dots\dots\dots (2)$$

may be used.

Hence

$$R\alpha\theta = \frac{s}{\frac{h}{R\theta} + \frac{3\theta}{2}}$$

and (1) may be written

$$s_\alpha = s \frac{\frac{h}{R\theta} - \frac{\theta}{2}}{\frac{h}{R\theta} + \frac{3\theta}{2}} \dots\dots\dots (3)$$

The mean power intercepted from the interfering rays is then  $2P_0(1 - \cos D) \frac{s_r}{s}$

and the excess is  $P_0 \frac{(s_\alpha - s_n)}{s}$ . The maximum and minimum values of  $P$  again occur where  $D = 2r\pi$  and  $D = (2r+1)\pi$ .

Then

$$\frac{P_{min}}{2P_0} = \frac{\theta}{\frac{h}{R\theta} + \frac{3\theta}{2}},$$

$$\frac{P_{max}}{2P_0} = \frac{2 \frac{h}{R\theta}}{\frac{h}{R\theta} + \frac{3\theta}{2}} \dots\dots\dots (4)$$

using the appropriate values of  $\theta$ .

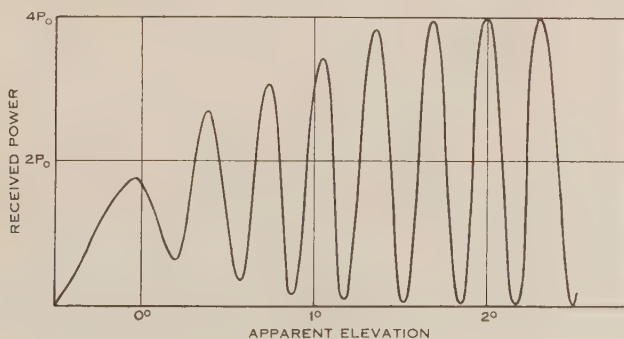


Fig. 9.—Theoretical interference pattern showing the effect of the earth's curvature on maxima and minima.



Figure 9 is a computed interference pattern based on these formulae, taking  $\lambda=3$  m.,  $h=300$  m., and neglecting the effect of receiver bandwidth. This pattern shows good agreement with actual patterns of the sources in Cygnus and Taurus obtained under the above conditions in New Zealand.

# GALACTIC RADIATION AT RADIO FREQUENCIES

## III. GALACTIC STRUCTURE

By J. G. BOLTON\* and K. C. WESTFOLD\*

[*Manuscript received November 4, 1950*]

### *Summary*

From radio-frequency observations it is deduced that the sun is situated in or near an arm of a spiral galaxy. The first part of this paper consists of an analysis of the radio-frequency data followed by the presentation of optical evidence in favour of a spiral form. The sense of rotation of the Galaxy—that of the spiral unwinding—is in accordance with the theories of Lindblad and Milne.

## I. INTRODUCTION

The discovery by Jansky(1) in 1932 of galactic noise at radio frequencies began a new era in astronomical investigation. However, it was not until 1940 that Reber(2) made observations at 160 Mc/s. with equipment specifically designed for such investigations. Since the development of highly sensitive receivers and directive aerial systems during the war, experimental progress has been more rapid and several surveys of noise distribution over the celestial sphere have been made.

So far, however, no satisfactory theory has been advanced as to the origin of galactic noise. Reber suggested a mechanism in terms of free-free transitions in interstellar matter and a quantum-mechanical treatment by Henyey and Keenan(3) showed good agreement with Reber's experimental results. Townes(4), however, has argued that the low-frequency (i.e.  $<100$  Mc/s.) results cannot be reconciled with Henyey and Keenan's formulae unless the electron temperature is of the order of  $150,000^{\circ}\text{K}$ . Various other hypotheses have been advanced. Pawsey, Payne-Scott, and McCready(5) suggested aggregate emissions from individual stars by processes similar to those operating on the sun at times of high sunspot activity. Greenstein, Henyey, and Keenan(6), however, pointed out that the large stellar dilution factor would make this mechanism highly improbable. More recently, Bolton and Stanley(7, 8) and Ryle and Smith(9) have shown that intense discrete sources make an appreciable contribution to the noise intensity from certain regions. However, the high concentration of noise towards the galactic plane, shown even in Jansky's results, leaves little doubt that the major part of the radiation originates within the Galaxy.

The present paper attempts to reconcile a number of noise contour maps at different frequencies, which together cover the whole celestial sphere. The effects of intense discrete sources are eliminated from these maps. In order

\* Division of Radiophysics, C.S.I.R.O., University Grounds, Sydney.

to interpret the noise distribution revealed by these surveys in terms of galactic structure, it is assumed that *the Galaxy is "optically thin" for the range of frequencies considered and that the material is more or less uniformly distributed.* The noise intensity in any direction is then taken as a measure of the spatial extent of the Galaxy in that direction.

## II. THE AVAILABLE DATA

### (a) *Published Surveys*

Table 1 contains details of the available plots of noise distribution over various parts of the celestial sphere between frequencies of 9.5 and 480 Mc/s.

TABLE 1  
SURVEYS OF GALACTIC NOISE

Survey	Author	Frequency (Mc/s.)	Beam Width	Date of Observations
1	Friis and Feldman(10)	9.5	$16^\circ \times 4^\circ$	1937
2	Jansky	18	$30^\circ \times 37^\circ$	1932
3	Friis and Feldman	18	$11^\circ \times 3^\circ$	1937
4	Franz(11)	30	$30^\circ$	1942
5	Moxon(12)	40	$35^\circ \times 70^\circ$	1946
6	Sander(13)	60	$20^\circ \times 30^\circ$	1946
7	Hey, Parsons, and Phillips(14)	64	$13^\circ \times 14^\circ$	1948
8	Moxon	90	$35^\circ$	1946
9	Bolton and Westfold(15)	100	$17^\circ$	1949
10	Reber	160	$12^\circ$	1940
11	Moxon	200		1946
12	Reber(16)	480	$3^\circ$	1948

The surveys 1 to 5 in this table are not considered suitable for the present investigation. The effect of ionospheric screening is not fully known for the lower frequencies and most of these surveys were made with aerials of low resolving power. Furthermore, the criterion of an optically thin galactic medium is probably not satisfied. Of the remaining surveys, 6, 8, and 11 were made with aerials of low resolving power leaving 7, 9, 10, and 12 suitable for detailed analysis.

The noise contours at 64 Mc/s., published by Hey, Parsons, and Phillips, and at 100 Mc/s., by the present authors, were derived by eliminating the effect of the aerial sensitivity pattern from the observed contours. This reduction was not attempted by Reber at 160 or 480 Mc/s., so that his contours are as seen by the aerial system. However, the beam widths of his aerial system were comparatively small. The areas covered by these four surveys are detailed in Table 2.

It can be seen from this table that the regions where the surveys overlap are within the ranges of galactic longitude,  $330$  to  $40^\circ$  and  $140$  to  $210^\circ$ . Where

TABLE 2  
AREAS COVERED BY VARIOUS SURVEYS

Author	Frequency (Mc/s.)	Limits of Declination	Approximate Limits of Galactic Longitude
Hey, Parsons, and Phillips	64	+50 to -30°	$\left\{ \begin{array}{l} 150 \text{ to } 210^\circ \\ 330 \text{ to } 50^\circ \end{array} \right\}$
Bolton and Westfold ..	100	+40 to -90°	140 to 40°
Reber .. .. .	160	+90 to -40°	330 to 200°
Reber .. .. .	480	+90 to -40°	320 to 200°

necessary, some reconciliation between the various surveys must, therefore, be attempted before a complete distribution can be constructed.

(b) *The Anticentre*

The region between galactic longitudes 140 and 220° which contains the anticentre of the Galaxy will first be considered. Figure 1 reproduces the intensity contours of the various surveys in their original form. The broken lines represent the boundaries of the surveys—a limit imposed by the latitude of the observing site. The 64 and 100 Mc/s. surveys show close general agreement except near longitude 210° where the 64 Mc/s. contours veer south of the galactic equator in the direction of increasing longitude. The 64 Mc/s. survey was probably affected by the low angle of observation, as is probable also in Reber's plot at 160 Mc/s. From the 0.5 contours of the 160 Mc/s. survey it would appear that the saddle-point occurs at longitude 205° and from the 1.0 contour at about 200°. The lower frequency plots in which aerial polar diagram effects have been eliminated favour a value of 198°, and this value will be accepted.

The great irregularity in the 480 Mc/s. results may be due to discrete sources in the constellation of Taurus(17) which have little influence at the lower frequencies (see Part I of this series).

(c) *The Galactic Centre and Cygnus*

Similar maps for this region are shown in Figure 2. There are marked differences between the various surveys near the galactic centre and in the constellation of Cygnus. A triple maximum between galactic longitudes 0 and 330° occurs in the 64 Mc/s. survey, a feature which is absent in the other three. Although it is possible that these maxima are features which are only apparent at frequencies lower than 100 Mc/s., observations of this region by the present authors using interference technique at 60 Mc/s. failed to find any trace of them. It seems that these maxima may be spurious, due perhaps to the circumstances of Hey's observations in this region.\*

\* In Hey's investigation each region was observed as it rose through the aerial beam in the east and again as it set in the west. In this way the effects of the side lobes of the aerial could be eliminated. The region near the centre of the Galaxy would be observed rising and setting with the aerial almost due south making analysis of the observations less reliable than for regions of more northerly declination.



In the constellation of Cygnus (longitude  $50^\circ$ ) Reber's survey at 480 Mc/s. shows a double maximum, whereas his survey at 160 Mc/s. and also that of Hey *et al.* show a single elongated maximum. The absence of a secondary peak on the lower frequencies is undoubtedly due to the low resolving power of the aerial

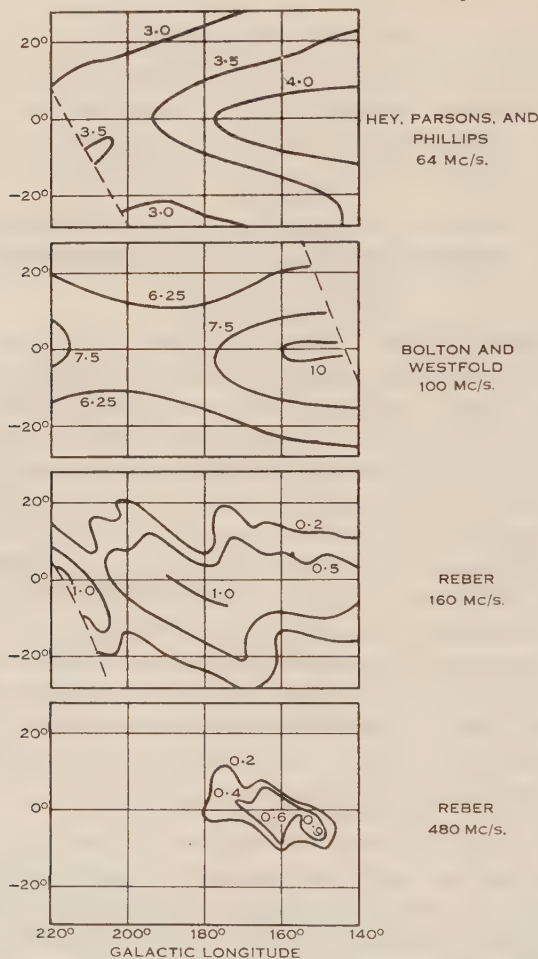


Fig. 1.—Galactic noise contours of various surveys between longitudes 140 and  $220^\circ$ .

systems employed. Figure 4 of Part I(15) shows this second peak as an interference pattern due to the discrete source in Cygnus, superimposed on a diffuse maximum. At 100 Mc/s. this discrete source contributes approximately 12 per cent. of the total radiation received from this region by an aerial of  $17^\circ$  beam width. The fraction at 64 Mc/s. would probably be slightly higher(18) (Part II of this series). The true secondary maximum in Cygnus is probably centred at galactic longitude  $50^\circ$  as shown in Reber's 480 Mc/s. survey.

*(d) Cassiopeia*

A further secondary maximum appears on both of Reber's surveys in the constellation of Cassiopeia (not illustrated). Ryle (personal communication) has informed the writers that the discrete source in Cassiopeia probably accounts entirely for this maximum.

From a comparison of Ryle's measurements(9) on the sources in Cassiopeia and Cygnus it is possible to correct Reber's contours at 160 Mc/s. in the Cygnus

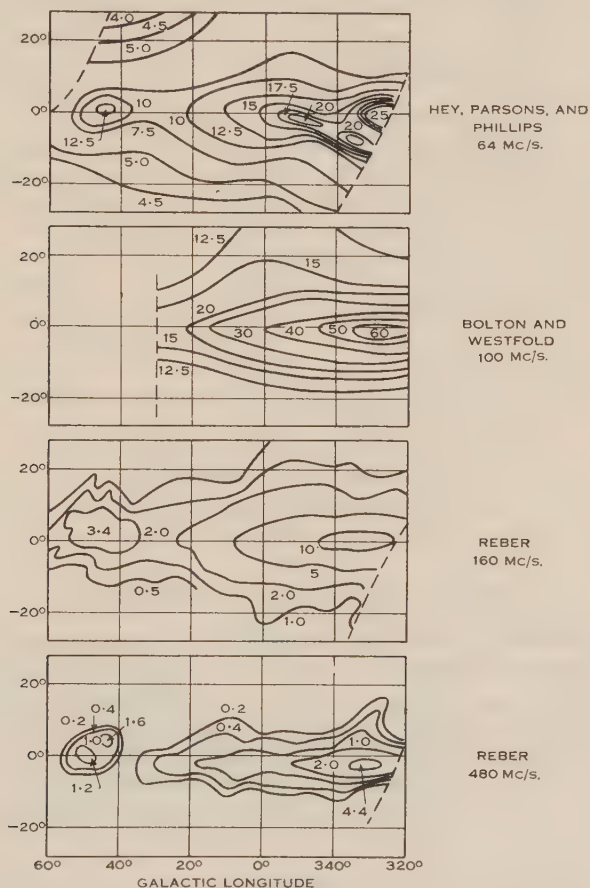


Fig. 2.—Galactic noise contours of various surveys between longitudes 320 and 60°.

region. At 80 Mc/s. the relative intensities from the sources in Cygnus and Cassiopeia are 14:23 for the "constant components" and 20:6 for the "fluctuating components". Weighting the fluctuating component by a likely factor of one-third, the total ratio becomes 8:10. Using this ratio (assuming that extrapolation to 160 Mc/s. is reasonable) and Reber's contours in Cassiopeia as an indication of aerial sensitivity pattern, the effect of the discrete source in Cygnus was eliminated, giving the result illustrated in Figure 3.

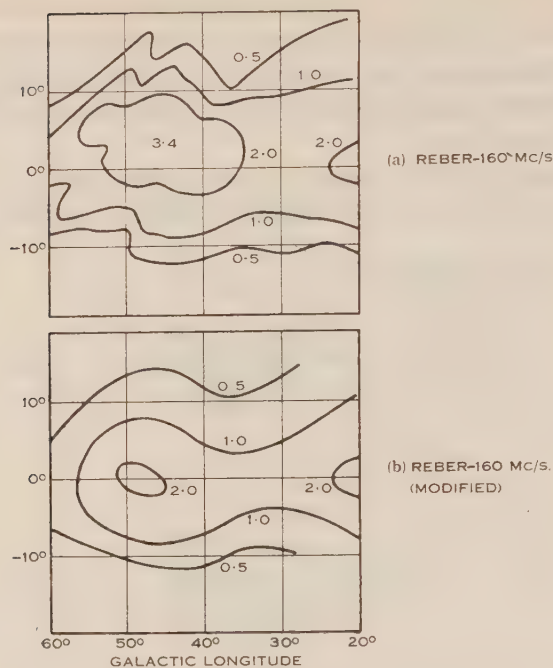


Fig. 3.—(a) Galactic noise contours at 160 Mc/s. for the Cygnus region.

(b) Probable form of contours at 160 Mc/s. without the discrete source.

### III. INTERPRETATION OF CONTOURS

It is now possible to examine the profiles of the contour maps along the galactic equator in the light of the previous discussion. These are shown in Figure 4, the scales being adjusted so that the ordinates near longitude 325° are approximately equal and convenient for comparison. The 160 Mc/s. results in the Cygnus region are taken from the modified contour map of Figure 3. Table 3 gives the ratios of intensity for the maximum in Cygnus and the galactic centre.

TABLE 3

Author	Frequency (Mc/s.)	Ratio of Intensities in Cygnus and Galactic Centre	Galactic Longitude of Maximum in Cygnus
Hey, Parsons, and Phillips ..	64	5 : 10	45°
Reber .. .. .	160	2.5 : 10	48°
Reber .. .. .	480	3 : 10	48°

Hey's value for the maximum in Cygnus requires some reduction to account for the discrete source. The ratio should probably be nearer to 4 : 10. Reber's results at 160 and 480 Mc/s. are affected by the sensitivity pattern of his aerial systems. As the maximum in Cygnus is sharper than that at the galactic centre, elimination of the aerial sensitivity pattern would increase Reber's ratios. A ratio of 3·8 : 10 will be taken to extrapolate the 100 Mc/s. profile beyond longitude  $20^\circ$ . This extrapolation\* is shown as the broken line in Figure 4. For general

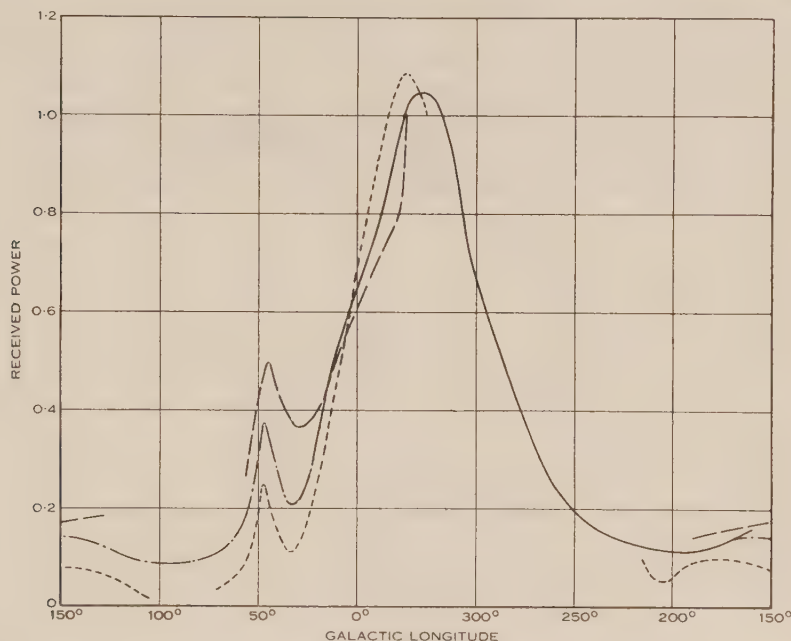


Fig. 4.—Contour profiles in the galactic plane (the scales are adjusted so as to afford easy comparison).

- — — Hey, Parsons, and Phillips 64 Mc/s.
- Reber 160 Mc/s.
- Bolton and Westfold 100 Mc/s.
- - — - Extension of 100 Mc/s. profile.

shape between longitudes  $20^\circ$  and  $40^\circ$ , Reber's contours in the modified form have been followed, as the minimum on Hey's plot is probably affected by the discrete source. Between longitudes  $60^\circ$  and  $140^\circ$  the available information is very sketchy but it is fairly clear that there is a shallow minimum in this region.

\* Although part of this region is visible from the latitude of the authors' observing site ( $34^\circ\text{S.}$ ) no measurements were made below  $15^\circ$  altitude (effectively above declination  $41^\circ\text{N.}$ ). Below this altitude part of the main lobe of the aerial intercepted the ground making results unreliable. Analysis of the observations was not taken beyond declination  $30^\circ\text{N.}$  However, some indication of the existence of the secondary maximum in Cygnus was gained from a subsidiary survey with the aerial directed horizontally over the sea towards the eastern horizon (see Part I(15)).



It appears that the general features of the noise distribution as shown by the contour profile in the galactic plane are a main maximum in the direction of the galactic centre ( $l=325^\circ$ ), a sharp secondary maximum towards the constellation of Cygnus ( $l=48^\circ$ ), and a small diffuse maximum\* probably centred between longitudes 150 and  $160^\circ$ .

In Figure 5, the 100 Mc/s. distribution in the equatorial plane is shown plotted in polar coordinates. On the assumption that the intensity in a given

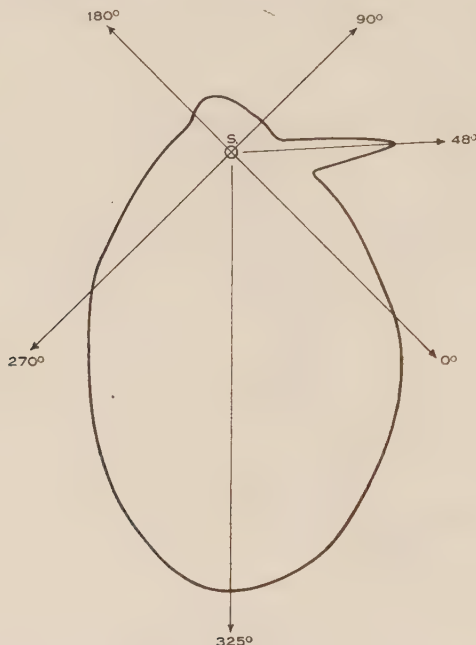


Fig. 5.—Variation of equivalent temperature at 100 Mc/s. in the galactic plane.

direction is a fair statistical measure of the extent of galactic material in that direction, this distribution can be interpreted as locating the sun in or near the arm of a spiral galaxy.† Such a distribution virtually limits the position of the sun to within the arm or between the arm and the nucleus, the arm running from Carina through the sun towards Cygnus. A derivation of the possible shape of the nucleus is given in Appendix I.

#### IV. ASTRONOMICAL CORRELATION

Since the development of the large reflector-type telescopes, progress in observation of extra-galactic nebulae has been rapid. Hubble has shown

\* The maximum centred slightly south of the galactic equator is in a region of conspicuous emission nebulae and contains at least three discrete sources.

† Although the authors' ideas followed the logical sequence outlined in this paper, it is worth noting that Reber(2) in 1944 suggested that the secondary maxima in Cygnus, Cassiopeia, and Perseus could represent extensions of a spiral galaxy.

that the various forms of these stellar systems fall naturally into a continuous sequence ranging from the globular or nearly spherical system through spheroidal systems of increasing eccentricity and simple spirals to very complex spirals. It is reasonable to believe that our own Galaxy falls into one of these broad classes, though absolute correspondence with any other system is not to be expected because most, at least of the spirals, have their own peculiarities in the distribution of dark matter, star clusters, etc.

It is generally recognized that the sun is situated at a considerable distance from the centre of the Galaxy. The estimated surface brightness of the Galaxy in the region of the sun is lower than that in parts of extra-galactic systems where structural features can easily be recognized from photographs. Further support is given by the great concentration of faint stars towards a region in the galactic plane round longitude  $325^\circ$ . Hubble has shown that the diameters of external systems increase steadily with increasing spiral structure; thus it is

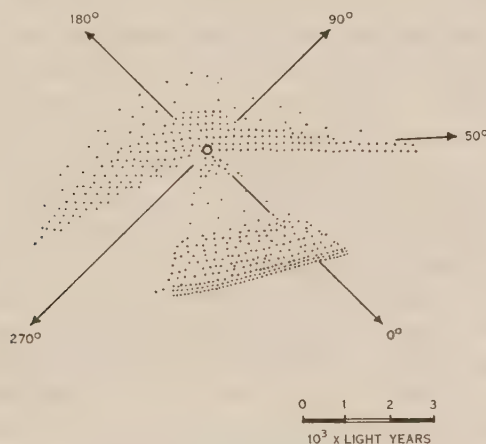


Fig. 6.—Pictorial representation of star distribution in the galactic plane near the sun.

likely that our own Galaxy, which compares with the largest of known systems, has some spiral features. It is difficult, however, to obtain direct evidence of any such features from observations made from the position of the sun. The star distribution in the neighbourhood of the sun is only known for certain regions up to a distance of 5000 light-years and results may be affected by an incomplete knowledge of the absorption in interstellar space. In general, however, present evidence lends some support to the spiral hypothesis.

In a recent summary, Bok states that about the anticentre ( $l$   $60$ – $170^\circ$ ) the star density falls off steadily with increasing distance from the sun. From Monoceros to Carina ( $l$   $170$ – $260^\circ$ ) the star density remains constant over a range increasing from 2000 to 4000 light-years, whilst towards Cygnus ( $l=50^\circ$ ) it is constant for up to 5000 light-years. For the southern half of the Milky Way, observations are by no means complete but between Aquila ( $l=10^\circ$ ) and

Sagittarius ( $l=340^\circ$ ) the star density shows an initial decrease with a minimum at 1500–2000 light-years followed by marked increases between 3000 and 5000 light-years.

Figure 6 is a pictorial representation of the probable situation in regions so far investigated. It seems that there is a concentration of stars towards a line running from Carina through the sun to Cygnus with a vacant lane between this line and a "nucleus". Even more marked is the concentration of *O* and *B* type stars towards this line. The average density of *O* and *B* stars decreases to 20 per cent. of the local density within 1000 light-years whereas, for all spectral types combined, the corresponding decrease is to 80 per cent. These figures include the Cygnus and Carina directions for which the densities of all types are constant up to 5000 light-years. *O* and *B* stars have only been observed in the spiral types of extra-galactic nebulae where they show preference for the arms of the spirals. The optical evidence is then consistent with the hypothesis of a spiral galaxy with an arm opening out towards Cygnus as suggested by the radio-frequency evidence. Further support might be found in a comparison of the dark nebulae forming the Rift and the dark lanes sometimes seen between the spiral arms and the nuclei in extra-galactic nebulae.

## V. GALACTIC ROTATION

### (a) *Direction of Rotation*

According to Smart, the hypothesis of galactic rotation was advanced as early as 1913 by Charlier and later in 1926 by Fotheringham and Schilt to interpret their observations on proper motions. There was, however, no striking evidence in its favour until Lindblad, in 1925, showed that it could account for the tendency of the velocity vectors of the stars of high velocity relative to the sun to avoid the hemisphere about longitude  $60^\circ$  in the galactic plane (approximately  $90^\circ$  from the galactic centre). Strömberg, in 1924, had made a thorough investigation of the velocities of four groups of stars. Each group showed a symmetrical dispersion of the velocity vectors about a point which could be interpreted as the mean velocity of the group relative to the sun. These four points lay on a line directed at longitude  $241^\circ$  in the galactic plane. Lindblad suggested that these four groups were subsystems of the whole galactic system, rotating with different angular velocities about an axis through the galactic centre and perpendicular to the galactic plane. The most flattened subsystem, the Milky Way clouds (including the sun), has the greatest angular velocity and the least flattened, the globular clusters, has the smallest. The observed asymmetry can be accounted for if the direction of rotation is *opposite to that of the increasing longitude*. Oort in 1927 developed Lindblad's ideas further, and was able to show that the double-wave variation with galactic longitude of the radial velocities, discovered by Freundlich and von der Pahlen in 1923, could be interpreted in terms of differential rotation, the stars nearer the galactic centre moving with greater angular velocity than those more distant. He was also able to interpret the results of an investigation on the asymmetry of the motions of the high-velocity stars in terms of the escape velocity from the system. Oort's hypothesis of differential rotation was confirmed by the detailed

observations of Plaskett and Pearce in 1936 on the radial velocities of *B* stars. At present it appears that these observations can be interpreted in terms of pure differential rotation of the Galaxy in which the velocity of the local group is 250 km./sec. in the direction of longitude  $55^\circ$  and the distance of the local system from the centre is about 8000 parsecs.

### (b) Sense of Rotation

It is evident that if the radio-frequency observations can be interpreted in terms of a spiral structure with an arm opening out in the direction of Cygnus, for our own Galaxy, at least, the sense of rotation is that of the spiral unwinding. The sense of rotation in extra-galactic systems may be determined only in certain special cases, where the nebulae are viewed almost edge-on. In these cases radial velocity measurements establish which end of the nebula is approaching and which receding with respect to the observer. The sense of rotation with respect to the spiral arms is dependent on the determination of the tilt of the fundamental plane, which cannot be directly observed but depends on the interpretation of the dark obscuring lanes. On the assumption that the obscuring material is highly concentrated towards the fundamental plane, Hubble(19) reports that of 15 nebulae which satisfy the required conditions, all appear to rotate in the same direction—with the arms trailing. Lindblad(20), however, finds that the observed intensity distribution in certain nebulae can be accounted for by the presence of a subsystem (see Section V (a)) of dark matter. The deduced tilt, and thus the sense of rotation, is then opposite to that claimed by Hubble, and in agreement with his theory of spiral structure in which the arms consist of matter ejected from unstable orbits in a rotating quasi-spheroidal system. Milne(21) deduces a similar structure from his kinematic theory of gravitation. The radio-frequency observations would appear to favour the theories of Lindblad and Milne.

## VI. ACKNOWLEDGMENT

The authors wish to thank Mr. Harley Wood, New South Wales Government Astronomer, for much helpful discussion of astronomical data.

## VII. REFERENCES\*

- (1) JANSKY, K. G.—Directional studies of atmospherics at high frequencies. *Proc. Inst. Radio Engrs. N.Y.* **30** : 1920-32 (1932).
- (2) REBER, G.—Cosmic static. *Astrophys. J.* **100** : 279 (1944).
- (3) HENYEV, L. G., and KEENAN, P. C.—Interstellar radiation from electrons and hydrogen atoms. *Astrophys. J.* **91** : 625-30 (1940).

\* In addition, reference was made to the following works :

- Bok, B. J., and Bok, P. F.—“The Milky Way.” (Blakiston, 1941.)  
 Bok, B. J.—“The Milky Way.” (Reprinted from Popular Astronomy, Vol. LII, 1944.)  
 Chandrasekhar, S.—“Principles of stellar dynamics.” (Chicago University Press, 1942.)  
 Smart, W. M.—“Stellar dynamics.” (Cambridge University Press, 1938.)  
 Plaskett, J. S., and Pearce, J. A.—“The motions of the O and B type stars and the scale of the Galaxy.” (Publications of the Dominion Astrophysical Observatory, Vol. V, No. 4.)



- (4) TOWNES, C. H.—Interpretation of radio radiation from the Milky Way. *Astrophys. J.* **105** : 235-40 (1947).
- (5) PAWSEY, J. L., PAYNE-SCOTT, RUBY, and MCCREADY, L. L.—Radio-frequency energy from the sun. *Nature* **157** : 158-9 (1946).
- (6) GREENSTEIN, J. L., HENYEV, L. G., and KEENAN, P. C.—Interstellar origin of cosmic radiation at radio frequencies. *Nature* **157** : 296 (1946).
- (7) BOLTON, J. G., and STANLEY, G. J.—A variable source of radio-frequency radiation in the constellation of Cygnus. *Nature* **161** : 312 (1948).
- (8) BOLTON, J. G.—Discrete sources of radio-frequency noise. *Nature* **162** : 141 (1948).
- (9) RYLE, M., and SMITH, F. G.—A new intense source of radio-frequency radiation in the constellation of Cassiopeia. *Nature* **162** : 462 (1938).
- (10) FRIIS, H. T., and FELDMAN, C. B.—A multiple unit steerable antenna for short-wave reception. *Proc. Inst. Radio Engrs. N.Y.* **25** : 841-917 (1937).
- (11) FRANZ, K.—Dimensionierung der Dreipunktgleichschaltung beim Gleichlauf von Überlagerungsempfängern. *Hochfrequenztech. u. Electroakust.* **59** : 143-50 (1942).
- (12) MOXON, L. A.—Variations of cosmic radiation with frequency. *Nature* **158** : 758-9 (1946).
- (13) SANDER, K. F.—Measurement of galactic noise at 60 Mc/s. *J. Instn. Elect. Engrs.* **93** (IIIA) : 1487-9 (1946).
- (14) HEY, J. S., PARSONS, S. J., and PHILLIPS, J. W.—An investigation of galactic radiation in the radio spectrum. *Proc. Roy. Soc. A* **192** : 425-45 (1948).
- (15) BOLTON, J. G., and WESTFOLD, K. C.—Galactic radiation at radio frequencies. I. 100 Mc/s. survey. *Aust. J. Sci. Res. A* **3** : 19 (1950).
- (16) REBER, G.—Cosmic static. *Proc. Inst. Radio Engrs. N.Y.* **36** : 1215-8 (1948).
- (17) BOLTON, J. G., and STANLEY, G. J.—The position and probable identification of the source of galactic radio-frequency radiation Taurus-A. *Aust. J. Sci. Res. A* **2** : 139-48 (1949).
- (18) STANLEY, G. J., and SLEE, O. B.—Galactic radiation at radio frequencies. II. The discrete sources. *Aust. J. Sci. Res. A* **3** : 234 (1950).
- (19) HUBBLE, E.—The direction of rotation in spiral nebulae. *Astrophys. J.* **97** : 112 (1943).
- (20) LINDBLAD, B.—On the dynamics of stellar systems. *Mon. Not. R. Astr. Soc.* **108** : 215 (1948).
- (21) MILNE, E. A.—On the spiral character of external galaxies. *Mon. Not. R. Astr. Soc.* **106** : 180-99 (1946).

## APPENDIX I

If the contours of equivalent temperature can be interpreted as giving contours of spatial extent as seen from the sun a possible shape for the nucleus of the Galaxy may be derived. The contour profiles in the meridian plane and the galactic plane will be used to construct two sections of a structure which would fit the radio-frequency observations. No importance, however, should be attached to the relative dimensions of the resulting ellipsoid, as there may well be a concentration of "noise sources" towards the centre of the Galaxy. Figure 7 is a contour profile of equivalent temperature in the meridian plane  $l=325^\circ$ , taken from the authors' 100 Mc/s. survey, and Figure 8 is a profile in the galactic plane, taken from the reconstructed 100 Mc/s. results in Figure 5. The hatched areas represent the portions which may reasonably be assigned to the nucleus so that the remainder\* represents "local structure". These results were used to construct the nuclear sections (Figs. 9 and 10). A ratio of 1:10 was assumed for the relative distances of the sun to the nearest point of the

\* It is probable that some of this noise is due to the integrated emissions from extra-galactic nebulae.

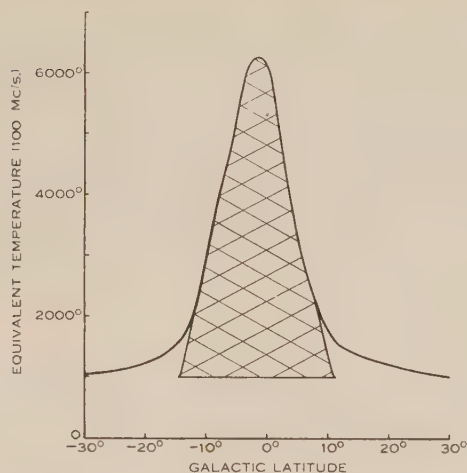


Fig. 7.—Contour profile of equivalent temperature at 100 Mc/s. in the meridian plane  $l=325^\circ$ .

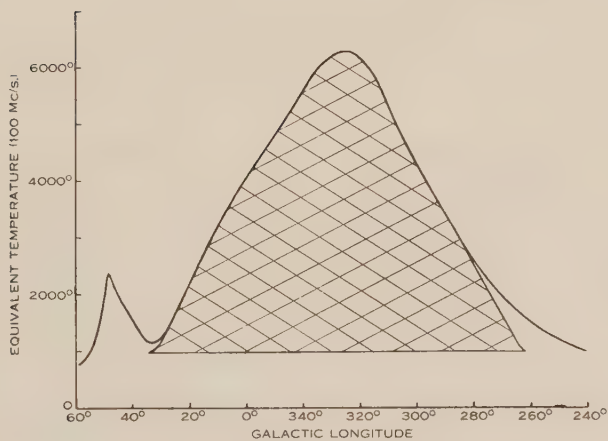


Fig. 8.—Contour profile of equivalent temperature at 100 Mc/s. in the galactic plane.

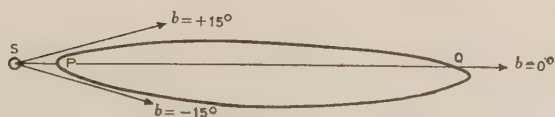


Fig. 9.—Meridian section of galactic nucleus at  $l=325^\circ$ .

nucleus and to the intercept through the centre\*; for example, the points  $P$  and  $Q$  in Figure 9 were determined by making  $PQ$  proportional to the ordinate of the curve in Figure 7 at galactic latitude  $b=0$  and  $SP : PQ$  as 1 : 10. Corresponding pairs of points were plotted from Figure 7, the points nearer  $S$  being judged so as to make the figure reasonably symmetrical. Figure 10, the section in the equatorial plane of the Galaxy, was constructed in a similar manner. Trial and error showed that the major axis must be inclined to the direction of longitude  $325^\circ$  so that its nearer end is towards the direction  $l=30^\circ$ . Points  $A$  and  $D$  in relation to  $S$  were determined in the same way as  $P$  and  $Q$  in Figure 9.

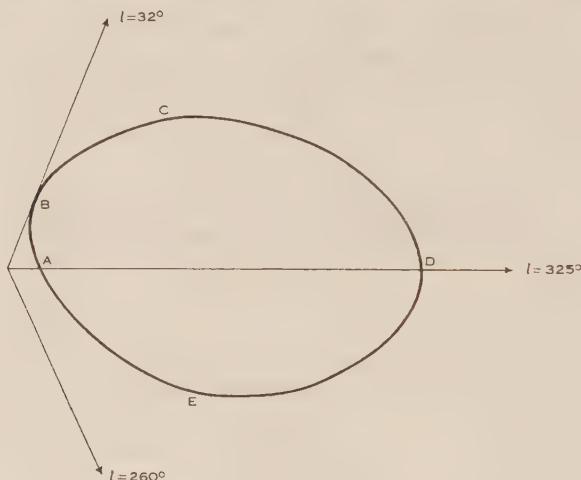


Fig. 10.—Equatorial section of galactic nucleus.

The distant points from  $D$  to  $C$  followed fairly simply, the placing of the near points in the direction of  $B$  making little difference to the general shape. From  $C$  to  $B$  the placing of the near and distant points was made so as to obtain a smooth curve. From  $D$  to  $E$  the curve was found to close smoothly up to the point  $E$  at  $l=290^\circ$  where points could no longer be fitted. It was therefore assumed that the near spiral arm was responsible for contributions from longitudes 290 to  $260^\circ$ , included in the hatched area. No attempt was made to discriminate another spiral arm which presumably is symmetrically situated. The ratios of the approximate ellipsoid thus obtained are 1 : 4.6 : 6.6.

\* This figure is in rough agreement with a distance of 8000 parsecs. from the sun to the galactic centre, and a distance of 5000 light-years to the region of increasing star density (see Section IV (b)).

# THE POLARIZATION OF THERMAL "SOLAR NOISE" AND A DETERMINATION OF THE SUN'S GENERAL MAGNETIC FIELD

By S. F. SMERD\*

[Manuscript received November 16, 1949]

## Summary

The equation of transfer of radiation and the magneto-ionic theory are used to derive expressions for the degree of polarization of thermal "solar noise" due to a general magnetic field of the sun. In particular, the net polarization of 600 Mc/s. (50 cm.) radiation corresponding to the maximum phase of the eclipse of November 1, 1948, as seen from Melbourne, Victoria, is evaluated theoretically and compared with observational evidence. This leads to an upper limit of 11 gauss for the surface field-strength at the solar poles at the time of observation.

## I. INTRODUCTION

Radio-frequency observations of the two circularly-polarized components of solar radiation taken at the time of an eclipse allow a deduction to be made of the magnitude of a possible general magnetic field of the sun. The determination of this field from other than optical observations seems desirable, particularly since some doubt has recently been thrown on the constancy, or even existence, of the sun's general magnetic field(1).

In this paper the polarization of thermal radio-frequency radiation due to a general magnetic field of the sun is obtained theoretically. The general magnetic field is represented by a central dipole in the plane of the sun's disk. The model of the solar atmosphere and the method of finding the intensity of radio-frequency radiation emerging from the solar atmosphere are those used in a previous paper(2). They are briefly summarized here in order to specify the assumptions on which the present treatment is based.

Finally, the method is applied to a particular case for which experimental data are available. The percentage circular polarization of 600 Mc/s. (50 cm.) radiation, corresponding to the maximum phase of the eclipse of November 1, 1948, as seen from Rockbank, near Melbourne, Victoria, is calculated. By comparing theoretical values with the observational evidence obtained by Christiansen, Yabsley, and Mills(3), an upper limit to the magnitude of the general magnetic field of the sun at the time of observation is evaluated.

## II. THE SOLAR ATMOSPHERE

The solar atmosphere is assumed to be a fully ionized hydrogen gas and spherically symmetrical. The chromosphere and corona are regarded as regions of uniform electron temperature. In numerical work we shall use a chromospheric temperature,  $T_{ch}$ , of  $3 \times 10^4$  °K. and a range of coronal temperatures,

\* Division of Radiophysics, C.S.I.R.O., University Grounds, Sydney.



$T_c$ , from  $2.5 \times 10^5$  to  $4 \times 10^6$  °K. The electron density in the chromosphere is represented by an exponential function. The exponential term is that given by Cillié and Menzel(4); the constant term has been adjusted to maintain continuity with the coronal density distribution at  $10^4$  km. above the photosphere. The corrected form of Baumbach's formula derived by Allen(5) is used for the electron density in the corona\*. This is

$$N = 10^8(1.55\rho^{-6} + 2.99\rho^{-16}), \dots\dots\dots (1)$$

where  $\rho$  is the distance from the centre of the sun in terms of its optical radius,  $R_0$ . The value of

$$R_0 = 6.95 \times 10^5 \text{ km.}$$

is used in numerical work.

The general magnetic field of the sun is represented by a central dipole in the plane of the sun's disk.

### III. THE INTENSITY OF RADIO-FREQUENCY RADIATION

#### (a) *The Emergent Intensity of a Ray*

The intensity of a ray emerging from the solar atmosphere is obtained by integration of the equation of transfer in the form given by Woolley(7), that is

$$\frac{d}{d\tau} \left( \frac{I}{\mu^2} \right) + \frac{I}{\mu^2} = B(T),$$

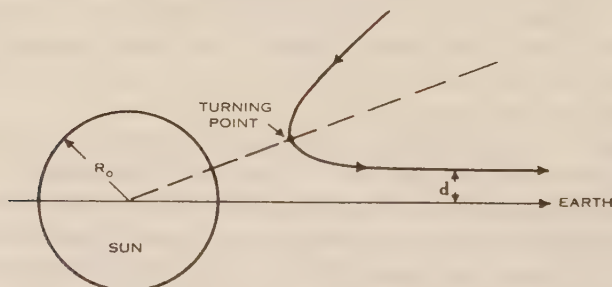


Fig. 1.—A typical ray trajectory from sun to earth.

where  $I$  is the intensity of radiation in the frequency range  $f, f+df$ ,

$\mu$  is the refractive index,

$B(T)$  is the intensity of unpolarized black-body radiation of frequency  $f$  and temperature  $T$ ,

$\tau$  is the optical depth, and is defined by

$$\tau = \int_s^\infty \kappa ds, \dots\dots\dots (2)$$

in which  $\kappa$  is the absorption coefficient, and

$ds$  is the path element along a trajectory.

A trajectory is specified by the parameter  $d$ , its distance from the sun-earth line where it emerges from the solar atmosphere (see Fig. 1). This parameter will be expressed in units of  $R_0$ .

\* A similar correction has also been given by v. d. Hulst(6).

The emergent intensity,  $I_e$ , of radiation along a trajectory is represented by the effective temperature,  $T_e$ , of the ray, the two quantities being related by

$$I_e = B(T_e). \quad (3)$$

We obtain (2) the following expressions for the effective temperature :

$$T_e(d) = T_c(1 - e^{-\tau_c(d)}) + T_{ch}e^{-\tau_c(d)}, \quad (4a)$$

for trajectories which penetrate into the chromosphere (subsequently referred to as "chromospheric trajectories"); and

$$T_e(d) = T_c(1 - e^{-2\tau_c(d)}) \quad (4b)$$

for trajectories whose turning points lie in the corona (subsequently referred to as "coronal trajectories"). Here,  $\tau_c(d)$  is the optical depth at the base of the corona in the case of chromospheric trajectories, and at the turning point in the case of coronal trajectories.

### (b) Ordinary and Extraordinary Rays

In the presence of a magnetic field, radiation in an ionized medium can propagate in two modes. These are conventionally referred to as the "ordinary" and "extraordinary" modes of propagation.

The physical properties of the solar atmosphere at any one frequency,  $f$ , are conveniently specified by the three dimensionless quantities  $x$ ,  $y$ , and  $z$ , commonly used in magneto-ionic theory. These quantities are defined by

$$\left. \begin{aligned} x &= f_0^2/f^2, \\ f_0^2 &= e^2 N / \pi m, \end{aligned} \right\} \quad (5)$$

in which

where  $f_0$  is the natural frequency of plasma oscillations,

$e$  is the electronic charge in e.s.u., and

$m$  is the electronic mass in grams;

$$\left. \begin{aligned} y &= f_H/f, \\ f_H &= eH/2\pi mc, \end{aligned} \right\} \quad (6)$$

in which

where  $f_H$  is the electron gyro-frequency,

$H$  is the magnetic-field intensity in gauss, and

$c$  is the free-space velocity of light in c.g.s. units;

and

$$z = \nu/2\pi f,$$

in which  $\nu$  is the frequency of elastic collisions of an electron with ions.

The quantity  $z$  is very much less than unity at all radio frequencies and in all parts of the solar atmosphere. The quantity  $y$  is less than unity at all frequencies greater than about 150 Mc/s. if the surface field-strength at the poles is not greater than 50 gauss. Radio-frequency propagation under such conditions has been discussed by Westfold,\* who gives the following approximations which apply when  $y$  is appreciably less than unity :

$$\left. \begin{aligned} \tau_o &= \tau - \tau', \\ \tau_e &= \tau + \tau'. \end{aligned} \right\} \quad (7)$$

\* Unpublished data.

The subscripts "o" and "e" denote the cases of "ordinary" and "extraordinary" radiation respectively. The optical depth,  $\tau$ , is given by (2), the absorption coefficient being that appropriate to the field-free case, namely (8)

$$\kappa = \frac{4e^6 A_1(2) N^2}{3\{2\pi(mkT)\}^{\frac{3}{2}} cf^2 \mu}, \quad \dots\dots\dots (8)$$

where  $k$  is Boltzmann's constant, and

$A_1(2)$  is a slowly varying function of  $N$  and  $T$  which may be taken as constant over large regions of the solar atmosphere.

The additive term in (7) is given by

$$\tau' = 2 \int_s^\infty \kappa y |\cos \Theta| ds, \quad \dots\dots\dots (9)$$

where  $\Theta$  is the angle between the magnetic field,  $H$ , and the direction of propagation.

A possible deviation of the helio-magnetic axis from the heliographic axis will be neglected. The latter is assumed to be perpendicular to the sun-earth line. The position of a point in the solar atmosphere will be specified by the spherical coordinates  $\rho$ ,  $\theta$ ,  $\phi$ . The distance from the centre of the sun,  $\rho$ , is in

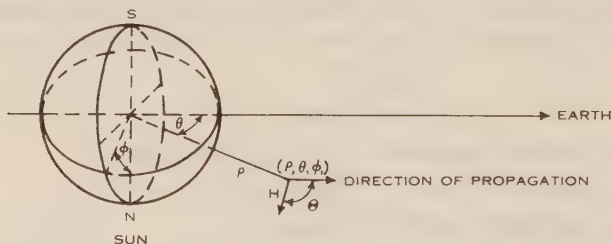


Fig. 2.—The system of polar coordinates to which points in the solar atmosphere are referred.

units of  $R_0$ ; the polar angle,  $\theta$ , is the angle between  $\rho$  and the sun-earth line; the azimuth,  $\phi$ , is in the plane of the sun's disk and measured from the heliographic north (see Fig. 2).

If  $M$  is the magnetic moment of the equivalent dipole representing the general magnetic field of the sun, the field strength at a point  $(\rho, \theta, \phi)$  is given by

$$H(\rho, \theta, \phi) = \frac{M}{R_0^3 \rho^3} (1 + 3 \sin^2 \theta \cos^2 \phi)^{\frac{1}{2}}; \quad \dots\dots\dots (10)$$

then from (6) and (10)

$$y = \left( \frac{e}{2\pi mc} \right) \frac{M}{R_0^3 \rho^3 f} (1 + 3 \sin^2 \theta \cos^2 \phi)^{\frac{1}{2}}, \quad \dots\dots\dots (11)$$

and for propagation parallel to the sun-earth line (that is, propagation in the non-deviating region, where the refractive index is unity for practical purposes) the angle  $\Theta$  is given by

$$\cos \Theta = \frac{3 \sin \theta \cos \theta \cos \phi}{(1 + 3 \sin^2 \theta \cos^2 \phi)^{\frac{1}{2}}}. \quad \dots\dots\dots (12)$$

The effective temperatures of ordinary and extraordinary rays, denoted by  $T_{e, o}$  and  $T_{e, e}$  respectively, are from (4) and (7)

$$\left. \begin{aligned} T_{e, o}(d) &= \frac{1}{2}[T_c - (T_c - T_{ch})e^{-\tau_c(d) + \tau'_c(d)}], \\ T_{e, e}(d) &= \frac{1}{2}[T_c - (T_c - T_{ch})e^{-\tau_c(d) - \tau'_c(d)}], \end{aligned} \right\} \dots\dots (13a)$$

in the case of chromospheric trajectories, and

$$\left. \begin{aligned} T_{e, o}(d) &= \frac{1}{2}T_c[1 - e^{-2\tau_c(d) + 2\tau'_c(d)}], \\ T_{e, e}(d) &= \frac{1}{2}T_c[1 - e^{-2\tau_c(d) - 2\tau'_c(d)}], \end{aligned} \right\} \dots\dots\dots (13b)$$

in the case of coronal trajectories.

In the case of the polarized radiation of ordinary and extraordinary rays the right-hand side of relation (3) is half the black-body intensity of the unpolarized radiation, thus explaining the factor  $\frac{1}{2}$  on the right-hand side of equations (13).

#### IV. THE POLARIZATION OF RADIO-FREQUENCY RADIATION FROM THE SUN

##### (a) *The Limiting Polarization*

The limiting polarization of solar radiation is defined as the polarization of radiation at emergence from the solar atmosphere. Under the conditions

$$x \ll 1,$$

$$y < 1,$$

the expressions given by Westfold\* for the complex polarization function,  $Q_{o, e}$ , reduce to

$$Q_{o, e} = \pm i \left( \frac{|\cos \Theta|}{\cos \Theta} - y \frac{\sin^2 \Theta}{2 \cos \Theta} \right). \dots\dots (14)$$

Since both  $x$  and  $y$  decrease outwards in the solar atmosphere to arbitrarily small values we obtain the limiting values of the complex polarization function—for  $\cos \Theta \neq 0$ —

$$Q_{o, e} = \pm i, \dots\dots\dots (15)$$

that is, the limiting polarization is circular. If right-handed polarization is taken to mean the clockwise rotation of the electric vector when viewed in the direction of propagation, the circular polarization of the ordinary ray will be left-handed and that of the extraordinary ray right-handed if  $\cos \Theta$  is positive.

For  $\cos \Theta = 0$ , that is, at the equator and at the limb of the radio-frequency disk, the limiting polarization of solar radiation will be linear. However, the rays showing linear polarization occupy a negligible fraction of the solar disk. Hence, if measurements of "solar noise" are carried out with two aerials accepting left-handed and right-handed circularly-polarized radiation respectively, these will measure the effective temperatures of ordinary and extraordinary rays.

##### (b) *The Degree of Polarization*

We define the degree of polarization of solar radiation along a trajectory by the function

$$p = \frac{T_{e, e} - T_{e, o}}{T_{e, e} + T_{e, o}}.$$

\* Unpublished data.



Hence from equations (13)

$$p = \frac{(1 - T_{ch}/T_c) \sinh \tau'_c}{e^{\tau_c} - (1 - T_{ch}/T_c) \cosh \tau'_c}, \dots\dots\dots (16a)$$

in the case of chromospheric trajectories, and

$$p = \frac{\sinh 2\tau'_c}{e^{2\tau_c} - \cosh 2\tau'_c}, \dots\dots\dots (16b)$$

in the case of coronal trajectories.

If the coronal part of the trajectory lies in the non-deviating region, the optical depth of the corona along a trajectory  $d$  in the field-free case is, from (1), (2), and (8)

$$\tau_c(d) = \frac{4e^6 A_1(2) R_o 10^{18}}{3c \{2\pi(mkT_c)^3\}^{\frac{1}{2}} f^2} \int_{\rho}^{\infty} \frac{2.40\rho^{-12} + 9.27\rho^{-22} + 8.94\rho^{-32}}{\sqrt{1 - d^2/\rho^2}} d\rho \dots\dots (17)$$

The solution of the integral is cumbersome and is omitted here.

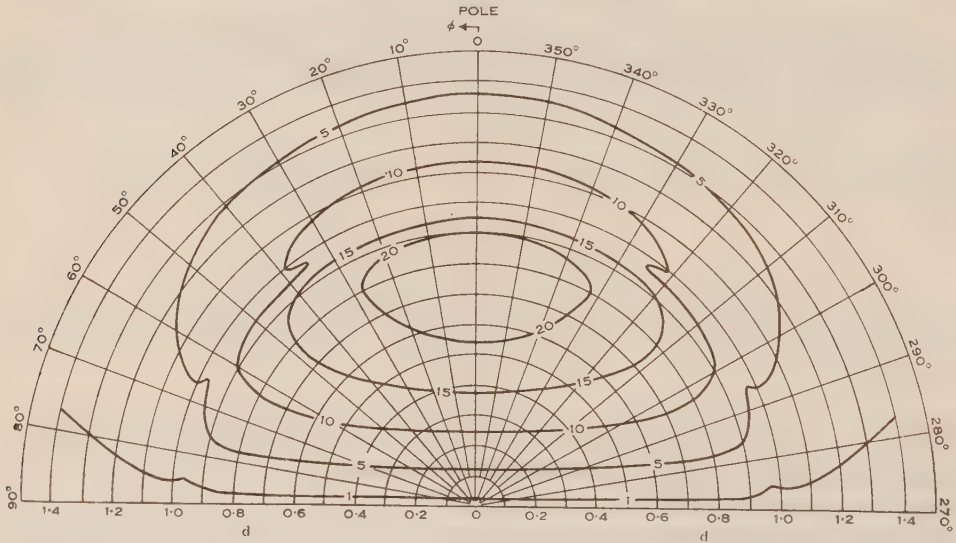


Fig. 3.—Contours of percentage circular polarization of 600 Mc/s. (50 cm.) solar radiation for a surface field-strength of 50 gauss at the poles. The distance from the centre of the sun is in units of the sun's optical radius. (The "kinks" in the contour lines are due to artificial discontinuities inherent in the treatment and are not significant.)

The value of  $\tau'$  over the coronal part of the trajectory ( $d, \varphi$ ) is from (9), (8), (1), (11), and (12)

$$\tau'_c(d, \varphi) = \frac{4e^7 A_1(2) 10^{16} M (0.16\rho^{-15} + 0.37\rho^{-25} + 0.26\rho^{-35})}{\{2m^5(\pi k T_c)^3\}^{\frac{1}{2}} c^2 f^3 R_o^2} d |\cos \varphi|, \dots\dots\dots (18)$$

where  $\rho$  is the lowest coronal height along the trajectory, that is, the height of the base of the corona in the case of chromospheric trajectories, and the height of the turning point in the case of coronal trajectories.

Figure 3 shows contours of percentage circular polarization of 600 Mc/s. (50 cm.) solar radiation calculated from (16), (17), and (18) for a million-degree

corona. The magnetic moment of the central dipole is taken to be  $8.4 \times 10^{18}$  gauss km.<sup>3</sup>, the corresponding surface field-strength at the poles being 50 gauss. The "kinks" in the contour lines are due to the discontinuity in the treatment of chromospheric and coronal trajectories, and no physical significance can be attached to them.

The net degree of polarization of solar radiation from a given section of the sun's disk is obtained by integrating  $p(d, \varphi)$  over the appropriate range of  $d$  and  $\varphi$ .

#### V. THE MAGNITUDE OF THE GENERAL MAGNETIC FIELD FROM ECLIPSE OBSERVATIONS OF THE POLARIZATION OF 600 Mc/s. SOLAR RADIATION

The eclipse of November 1, 1948, at its maximum phase as seen from Rockbank, near Melbourne, is illustrated in Figure 4. Figure 5 shows the calculated percentage circular polarization of 600 Mc/s. (50 cm.) radiation corresponding

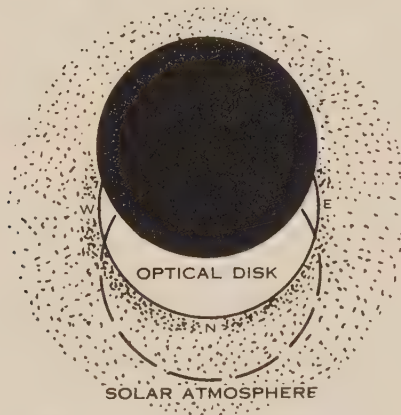


Fig. 4.—The maximum phase of the eclipse of November 1, 1948, as seen from Rockbank, near Melbourne, Victoria. The dotted line encloses that portion of the radio-frequency disk which contributes to the net polarization of radiation. (Outside this line the polarization of rays in the southern hemisphere is balanced by the opposite polarization of the corresponding rays in the northern hemisphere.)

to the phase depicted in Figure 4 for coronal temperatures\* of between  $2.5 \times 10^5$  and  $4 \times 10^6$  °K. The magnetic moment of the central dipole, representing the general magnetic field of the sun, is again assumed to be  $8.4 \times 10^{18}$  gauss km.<sup>3</sup>. The polarization shows a marked dependence on the coronal temperature.

Within the limits of observational accuracy (quoted as 3 per cent. predominance of one circularly-polarized component), Christiansen, Yabsley, and Mills(3) did not detect any polarization at the maximum phase of the eclipse. An upper limit of 1.5 per cent. is therefore placed on the observed value of

\* Throughout this discussion the coronal temperature refers to a mean value of the electron temperature in the coronal region of origin of 600 Mc/s. radiation.

circular polarization. This value is lower than the lowest theoretical value derived above. In order to obtain sufficiently low theoretical values it is necessary to assume

- (i) coronal temperatures less than  $2 \times 10^5$  °K. or considerably greater than  $10^7$  °K. (values obtained by extrapolation of Fig. 5), or
- (ii) magnetic fields considerably smaller than that initially assumed.

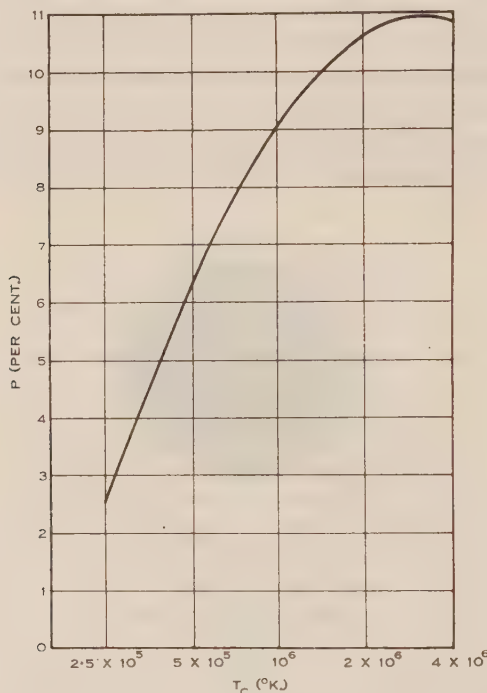


Fig. 5.—The variation of the percentage circular polarization of 600 Mc/s. (50 cm.) solar radiation (corresponding to the maximum phase of the eclipse of November 1, 1948) with coronal temperature for a surface field-strength of 50 gauss at the poles.

Now in the absence of a knowledge of the intensity distribution across the 600 Mc/s. solar disk a unique value of the coronal temperature cannot be derived from the available observational evidence (see Smerd 2). A comparison, however, of observed and theoretical values(2) of the “apparent temperature” (an equivalent measure of the flux density at the earth) of the sun at 600 Mc/s., makes it appear improbable that the coronal temperature at the time of the eclipse would have been less than about  $5 \times 10^5$  °K. or greater than about  $2 \times 10^6$  °K. Thus, the first possibility, ((i) above) may be discarded.

It remains then to calculate the magnetic field in the corona, in terms of the magnetic moment of an equivalent dipole at the centre of the sun, which

would result in the observed upper limit of 1.5 per cent. of the percentage circular polarization. This has been done for coronal temperatures of

- (i)  $5 \times 10^5$  and  $2 \times 10^6$  °K., that is, the lower and the upper limits suggested by the theoretical interpretation of the apparent temperature, and
- (ii) the conventional value of  $10^6$  °K.

The results of these calculations are collected in Table 1 below.

TABLE 1  
THE UPPER LIMITS OF THE SUN'S GENERAL MAGNETIC FIELD DERIVED  
FROM ECLIPSE OBSERVATIONS (NOVEMBER 1, 1948)

Coronal Temperature $T_c$ ( $10^6$ °K.)	Upper Limit of:	
	Magnetic Moment of Central Dipole $M$ ( $10^{18}$ gauss km. <sup>3</sup> )	Surface Field- Strength at the Poles $H$ (gauss)
0.5	1.8	10.7
1.0	1.4	8.2
2.0	1.2	7.1

We conclude, therefore, that the magnetic field in the coronal part of the region of origin of 600 Mc/s. (50 cm.) radiation at the time of the eclipse was certainly not greater than that due to a central dipole of moment  $2 \times 10^{18}$  gauss km.<sup>3</sup>. The values of surface field-strength listed in Table 1 are calculated from the corresponding values of the dipole moment by equation (10), and so do not include the effects of a possible "magnetic shielding" of the corona from the sun. Low values (considerably less than 50 gauss) of the surface field-strength at the poles have, however, also been obtained at different times by Babcock(1). These values are based on optical measurements which refer to near-surface levels of the sun (photosphere and reversing layer).

## VI. REFERENCES

- (1) BABCOCK, H. D.—*Publ. Astr. Soc. Pacif.* **60**: 244 (1948).
- (2) SMERD, S. F.—*Aust. J. Sci. Res. A* **3**: 34 (1950).
- (3) CHRISTIANSEN, W. N., YABSLEY, D. E., and MILLS, B. Y.—*Aust. J. Sci. Res. A* **2**: 506 (1949).
- (4) CILLIÉ, G. G., and MENZEL, D. H.—*Harv. Coll. Obs. Circ. No.* 410 (1935).
- (5) ALLEN, C. W.—*Mon. Not. R. Astr. Soc.* **107**: 426 (1947).
- (6) V. D. HULST, H. C.—*Astrophys. J.* **105**: 471 (1947).
- (7) WOOLLEY, R. v. d. R.—*Suppl. Aust. J. Sci.* **10**(2): i (1947).
- (8) SMERD, S. F., and WESTFOLD, K. C.—*Phil. Mag.* **40**: 831 (1949).



# RADIATIVE HEAT TRANSFER IN THE AIR NEAR THE GROUND

By E. L. DEACON\*

[Manuscript received November 30, 1949]

## Summary

A radiation chart for use in calculating the rates of heating or cooling by long wave radiation of air layers in the lowest 100 m. of the atmosphere, is presented. It takes account of the radiation due to carbon dioxide in a more satisfactory manner than does Elsasser's chart which applies to the relatively deep atmospheric layers dealt with in synoptic problems.

As an example of the application of the chart, the rate of radiative cooling of the lowest 80 m. of the atmosphere, just before dawn, is evaluated from published data for clear June nights of little wind in southern England. It is shown that, under these conditions, radiative exchanges in the air play a large part in cooling the air in the surface inversion layers.

## I. INTRODUCTION

The study of the transfer of sensible heat to the atmosphere from the underlying surface requires a knowledge of the processes of heat transfer by long wave radiation as well as by turbulent convection and molecular conduction. In the troposphere long wave radiation exchanges depend almost entirely on the characteristics of the infra-red absorption spectra of water vapour and carbon dioxide, as the simple gases in the atmosphere play no appreciable part, being almost completely transparent to long wave radiation as to the short wave radiation from the sun.

The classical investigations of Simpson(1, 2) and Brunt(3, 4) on the role of long wave radiation in the atmosphere were handicapped by the lack of reliable quantitative data on the intensity of absorption of long wave radiation by the water vapour in the atmosphere; subsequent work has shown that the absorption indicated by Hettner's work on steam, which was the only guide at that time, is very much greater than occurs with water vapour under atmospheric conditions.

In recent years, the advancement of knowledge in this field enabled Elsasser(5) to give a more adequate treatment of the long wave radiation problem and to develop a radiation chart from which the radiative heating or cooling of tropospheric air layers may be calculated with reasonable accuracy, given the distribution with height of temperature and water vapour. His method takes account of the variations in absorption by water vapour caused by the considerable ranges of water path,† pressure, and temperature encountered in the

\* Section of Meteorological Physics, C.S.I.R.O., Melbourne.

† In the following, water path or carbon dioxide path between two levels is the amount of the substance in question in a vertical column of unit cross-section area extending between those levels and is accordingly given in units of g. cm.<sup>-2</sup>.

troposphere. This, in itself, is a problem of considerable complexity and to have treated the carbon dioxide absorption in an equally thorough manner would have introduced formidable complications in view of the wide range of variation of the water vapour : carbon dioxide ratio from level to level in the atmosphere. As Elsasser's object was the development of a treatment simple enough for ready application to synoptic problems and as, in that case, the emphasis is on the heating or cooling of relatively thick atmospheric layers (hundreds of metres thick), he was able to make the simplifying assumption that even thin layers (in the synoptic sense) are completely opaque in the spectral region occupied by the carbon dioxide absorption band.

When the heating or cooling of the lowest few metres or decametres of the atmosphere is required, then Elsasser's treatment of the carbon dioxide absorption is not sufficiently accurate. In the following, a radiation chart more appropriate to micrometeorological use is presented together with an example of its application to a specific problem—the evaluation of the rate of cooling by radiation of the lowest 80 m. of the atmosphere during the hour or two before dawn on clear summer nights in England. Selection of this example is determined by the present availability of suitable published material.

## II. THE RADIATION CHART

The experimental values for the emissivity of carbon dioxide in air (to the extent of 0.03 per cent. by volume) indicate the following variation with path length (Elsasser 5):

Thickness of air layer (m.)	..	..	..	25	100	400	1600
Emissivity of CO <sub>2</sub> (% of black body)	..	..	..	3.7	8.2	12.6	16.7

The assumption made by Elsasser, in constructing his radiation chart, that even thin layers of the atmosphere are opaque in the spectral region occupied by the carbon dioxide absorption band, would, if strictly true, entail an emissivity constant with path length. The above figures show that while the emissivity only increases by 30 per cent. with increase of thickness from 400 to 1600 m., it more than doubles with increase from 25 to 100 m. thickness. It is apparent, therefore, that Elsasser's assumption is not a good approximation for use in considering the lowest 100 m. of the atmosphere. In this layer it is a closer approximation to consider the carbon dioxide absorption to be additive to that due to water vapour as the absorption bands of the two substances overlap only to a small extent. For path lengths (through air) of the order of 5 m. the error introduced is negligible(6), but increases with the greater path lengths, the absorption being over-estimated. It will be seen later, however, that the error is unimportant in treating the lowest hundred metres of the atmosphere.

In constructing a radiation chart for application to micrometeorological problems, a simplification can be introduced as a result of a peculiarity of the water vapour infra-red absorption spectrum in the relatively restricted range of temperatures encountered in these studies. Temperature affects the radiation absorption by water vapour in two ways, firstly by shifting the black-body energy distribution relative to the wavelengths of strong absorption by water vapour

and secondly by changing the intensity of the absorption bands. The first effect leads to a reduction in absorption with increasing temperature due to the progressive shift of the peak of the black-body energy distribution away from the intense rotational absorption band of the water vapour spectrum towards the transparent region around  $10\mu$  wavelength. On this account, by itself, the absorptivity at  $20^\circ\text{C}$ . would be from 15 to 35 per cent. less than at  $-40^\circ\text{C}$ ., the greater variation being for the smaller water paths. The second effect, the increasing intensity of the absorption bands (mainly in the wings of the bands) with rising temperature, works in opposition to the first and, although the experimental evidence is meagre, it appears from Elsasser's work to be of similar magnitude. In Table 1 emissivities (or absorptivities) of water vapour (for diffuse radiation) have been calculated from data given by Elsasser(5, Table 8).

TABLE 1  
TEMPERATURE VARIATION OF WATER VAPOUR EMISSIVITY

Water Path (g. cm. <sup>-2</sup> )	Emissivity (% Black Body) at		
	$-40^\circ\text{C}$ .	$-10^\circ\text{C}$ .	$20^\circ\text{C}$ .
0.00025 .. ..	6.1	5.6	5.3
0.001 .. ..	12.8	11.7	11.1
0.01 .. ..	26.8	26.0	26.3
0.1 .. ..	44.5	45.1	47.1

It is apparent from these values that over this range of variables the two temperature effects nearly balance and, as water vapour emissivity measurements have not yet attained a high degree of precision, the small resultant variation may be neglected. Emissivity values independent of temperature will therefore be used for the present purpose.

With carbon dioxide the main absorption band operative in radiation transfer in the lower atmosphere is that centred at  $15\mu$ . In this case the change in emissivity with temperature is almost entirely due to changing intensity and width of the absorption band, the effect due to displacement of the black-body distribution relative to the band being very small. Strictly, therefore, an emissivity of carbon dioxide increasing somewhat with temperature should be employed but, in practice, a constant emissivity will suffice as the carbon dioxide absorption is of secondary importance to that of water vapour in the lower atmosphere.

The variation of radiation absorption by water vapour and carbon dioxide with changing atmospheric pressure is also negligible in the lowest hundred metres or so of the atmosphere, the absorption being approximately proportional to the square root of the pressure.



Emissivities independent of temperature enable a radiation chart to be constructed directly from the emissivities of moist air and carbon dioxide which have been determined by laboratory experiments. The chart is then based on the following relationship for the diffuse radiation from an isothermal atmospheric layer of depth  $d$ :

$$\text{radiation flux} = e_d \sigma T^4,$$

where  $e_d$  = total emissivity of water vapour and carbon dioxide in an atmospheric layer of depth  $d$ ,

$\sigma$  = Stefan's constant,

$T$  = absolute temperature of air layer.

The abscissae of the chart are made proportional to  $T^4$  and the ordinates to  $(1 - e_d)$ . The emissivities of water vapour and carbon dioxide for diffuse radiation (called *slab emissivities*) needed to evaluate  $e_d$  are obtained from the measured emissivities for unidirectional radiation (*column emissivities*) using relationships given by Elsasser(5, p. 35). Over the range of path length ( $u$ ) for which absorption is proportional to  $\sqrt{u}$ , the slab emissivity of a layer of depth  $d$  is equal to the emissivity for unidirectional radiation of a column of length  $1.78d$ . With increasing  $u$  above this range the corresponding factor tends to 1.5 while for very small values of  $u$  the factor is somewhat greater than 1.78.

The following water vapour emissivities at 20 °C. have been used in the construction of the chart. They were taken from a fair curve drawn through Brooks's(6) values from laboratory experiments up to 0.001 g. cm. water path and through Elsasser's(7) experimental values for the greater paths to which Brooks's measurements did not extend.

Water path (g. cm. <sup>-2</sup> )	..	0.0001	0.0004	0.001	0.004	0.01	0.04	0.1
Emissivity (column)	(%)							
black body)	..	1.6	4.9	8.6	15.3	20.5	30	37

For carbon dioxide the emissivity values given by Elsasser(5, Figs. 26 and 27) have been used. They are based on the results of several experimenters which are in good concordance.

For a given water path in the atmosphere, the associated carbon dioxide path is inversely proportional to the water vapour pressure, provided the carbon dioxide concentration in the atmosphere is constant. In constructing the chart, carbon dioxide is taken to constitute 0.03 per cent. by volume of the atmosphere(8) and assuming, as mentioned previously, that the emissivities of water vapour and carbon dioxide are directly additive, four sets of water path lines have been inserted on the chart (Fig. 1) corresponding to water vapour pressures of 5, 10, 20, and 40 mb.

A specimen of the radiation chart is shown in Figure 1 while in Table 2 below the ordinates of the water path lines are given so that the chart (or a portion of it) may be redrawn to any scale found convenient by the individual user. The properties of a radiation chart of this type are given in considerable detail by Elsasser(5, pp. 20-27) to which paper reference should be made for a



full account. An example given in the following shows the way the chart may be used to estimate the heating or cooling of various air layers near the ground through the agency of radiation exchanges.

TABLE 2  
ORDINATES OF THE RADIATION CHART

Water Path (g. cm. <sup>-2</sup> )	Ordinates for Vapour Pressures of			
	5 mb.	10 mb.	20 mb.	40 mb.
0	100	100	100	100
0.0001	97.1	97.2	97.2	97.2
0.0002	94.9	95.1	95.2	95.2
0.0004	92.0	92.3	92.4	92.5
0.0006	90.0	90.3	90.5	90.6
0.0008	88.6	89.0	89.3	89.4
0.001	87.5	87.9	88.2	88.3
0.002	83.5	84.2	84.7	85.0
0.004	79.1	80.1	80.9	81.5
0.006	76.0	77.7	78.5	79.0
0.008	73.6	75.5	76.6	77.1
0.01	71.7	73.8	75.0	75.7
0.02	64.8	67.1	69.0	70.5
0.04	57.2	59.6	62.0	64.0
0.06	52.6	55.0	57.3	59.6
0.08	49.2	51.7	54.1	56.4
0.1	47.0	49.1	51.4	53.6
0.2	38.9	40.1	43.0	45.6
∞	0	0	0	0

Values beyond 0.2 g. cm.<sup>-2</sup> water path are not included in Table 2 as for such water paths the error due to treating the emissivities of water vapour and carbon dioxide additively would become appreciable. The extent of the error up to 0.2 g. cm.<sup>-2</sup> is only of similar magnitude to the uncertainties in the experimental data for atmospheric absorption as may be seen from the fact that Brooks's experimental curve(6, Fig. 5) based on sky observations of atmospheric radiation at about 4 mb. vapour pressure indicates a combined emissivity of water vapour and carbon dioxide of 58 per cent. black body as compared with a value of 61 per cent. arrived at by the method used in constructing Table 2.

### III. RADIATIVE COOLING OF THE LOWEST 80 METRES OF THE ATMOSPHERE ON CLEAR NIGHTS WITH LITTLE WIND

Measurements of air temperature at five heights between 1.2 and 88 m. over open country at Leafield, Oxfordshire, have been published by Johnson and Heywood(9). For the bulk of these observations ground surface temperatures

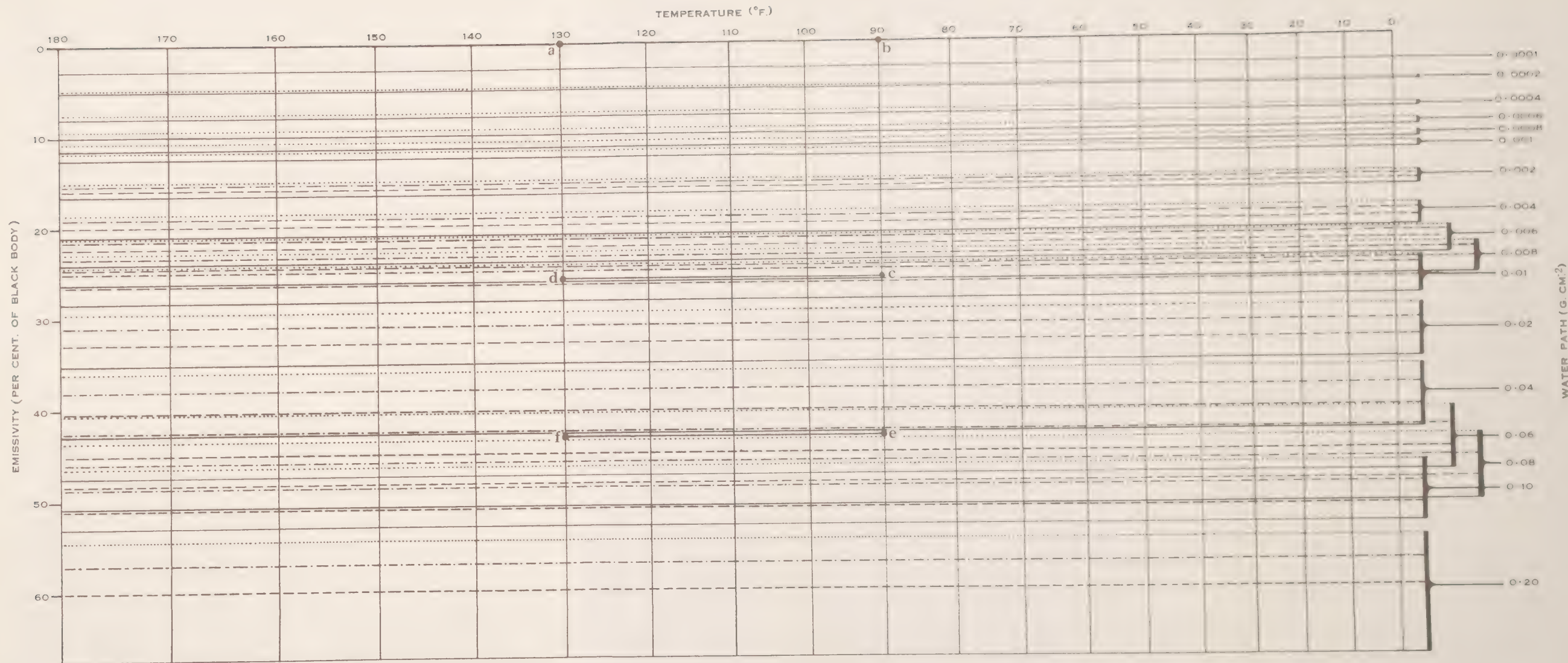


Fig. 1.—Radiation chart. Area of rectangle abef =  $0.1 \text{ cal. cm}^{-2} \text{ min}^{-1}$ ; area abcd =  $1 \text{ mcal. cm}^{-2} \text{ sec}^{-1}$   
 Water vapour pressure: ——— 5 mb.; - - - - 10 mb.; - · - · - 20 mb.; ····· 40 mb.



are not available. In Table 28 of their paper, however, the mean temperature profiles for a period just before dawn on clear June nights with little wind is given together with the mean grass minimum temperature. Robinson(10) has found from radiation measurements at Kew that the reading of a terrestrial radiation thermometer gives the radiative temperature of a grass surface to a good approximation. The grass minimum temperature is therefore taken to be the appropriate surface temperature to use with the air temperature profile at 0400 G.M.T., which is approximately the time of minimum temperature in June.

Due to the non-linear relationship between temperature and the saturation vapour pressure of water, the mean vapour pressure at a height of 1.2 m. given in the above-mentioned table, indicates supersaturation at the mean air temperature given for this height. In the following it is assumed that, at each height, the vapour pressure was the saturation vapour pressure corresponding to the temperature at that height—a reasonable assumption for the conditions of the observations. The calculation of water paths for use with the radiation chart to find the radiational cooling of the 0 to 10 m. air layer, is shown in Table 3.\*

TABLE 3

WATER PATHS FOR CLEAR JUNE NIGHTS AT LEAFIELD FROM DATA OF JOHNSON AND HEYWOOD(9)

Height, $z$ (m.)	Temp. (°F.)	Vapour Pressure (mb.)	Mean Vapour Pressure From 0 to $z$	Water Paths (g. cm. <sup>-2</sup> )	
				For Level $z=0$	For $z=10$ m.
0	33	6.4	—	0	0.0081
1.2	44.5	10.0	9	0.00085	0.0073
10	48.1	11.4	10.5	0.0081	0
20	49.1	11.9	11	0.017	0.009
40	50.1	12.3	11.5	0.035	0.027
60	50.7	12.6	12	0.055	0.047
80	51.0	12.8	12	0.074	0.066

The water paths given in the fifth column above are the amounts of water vapour in a vertical column of cross section 1 cm.<sup>2</sup> extending from the surface up to the level in the first column. Those in the last column are water paths upward and downward from the 10 m. level. The temperatures (column 2) are now plotted on the radiation chart (see Fig. 2) at the appropriate water path values. In this example it will be sufficiently accurate to plot the points with

\* From a curve drawn through the experimental values of Table 28 of Johnson and Heywood(9), the temperatures at more convenient heights have been interpolated.



respect to the set of water path lines for 10 mb. vapour pressure. The curve *ABC* is drawn through the points for the water path values in the fifth column above, while using the water paths in the last column gives a curve with two branches *GKL* and *GHJ*.

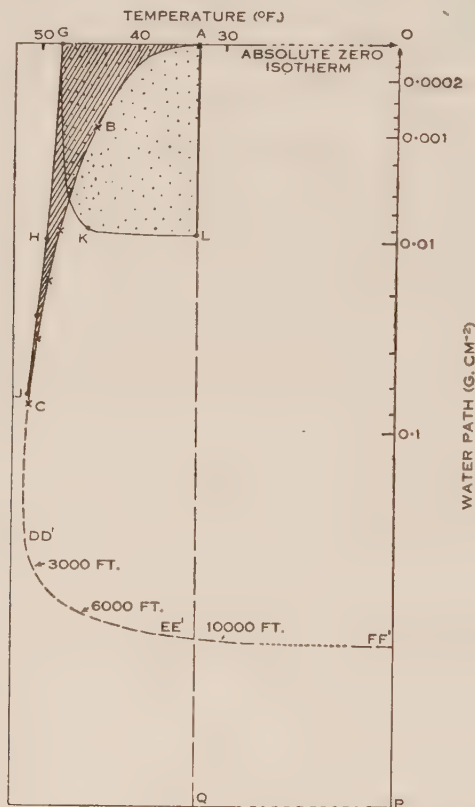


Fig. 2.—Radiation chart plotting for 0400 G.M.T. on clear June nights at Leafield, Oxfordshire.

It will be noted that the two curves *ABC* and *GHJ* converge and nearly meet at the points *J* and *C* which correspond to the values at 80 m. height. Had the temperature and humidity structure of the atmosphere been observed up to great heights in the atmosphere, the curve *ABC* could be continued along some such path as *CDE* to meet the absolute zero isotherm at *F*. (The extension shown in Figure 2 is based on upper air data for similar weather situations to those tabulated.) The similar extension of the curve *GHJ* would coincide with *CD* over the isothermal part of the ascent curve and then cross over and run slightly above *DEF* (through points referred to as *D'E'F'* in the following) for the greater heights where lapse conditions exist. The separation of the two curves *DEF* and *D'E'F'* is so small on the scale of Figure 2 as to be incapable of representation.

As shown by Elsasser(5, pp. 24-27) the following long wave radiation fluxes are represented by areas on the chart (see Fig. 2).

$$\begin{array}{lcl}
 \text{Downward fluxes} & \left\{ \begin{array}{l} \text{at ground level from} \\ \text{whole atmosphere} \end{array} \right. & = \text{Area } ABCDEFOA \\
 & \left\{ \begin{array}{l} \text{at 10 m. from atmos-} \\ \text{phere above} \end{array} \right. & = \text{Area } GHJD'E'F'OG \\
 & \text{Difference} & = \text{Area } GHJD'E'F'FEDCBAG \\
 & & \text{(shown cross-hatched)} \\
 \\ 
 \text{Upward fluxes} & \left\{ \begin{array}{l} \text{from ground (assumed} \\ \text{to radiate as a black} \\ \text{body)} \end{array} \right. & = \text{Area } AOPQA \\
 & \left\{ \begin{array}{l} \text{from top of 10 m. layer} \end{array} \right. & = \text{Area } QLKGOPQ \\
 & \text{Difference} & = \text{Area } ALKG \\
 & & \text{(shown stippled)}
 \end{array}$$

The difference between the stippled and cross hatched areas in Figure 2 accordingly represents the net rate of loss of heat of the 0 to 10 m. layer and this may conveniently be found by running the point of the planimeter round the course *ALKGHJCBA*.\* This neglects the very small area between *DEF* and *D'E'F'* which should strictly be taken into account but in the present example it is estimated using Elsasser's chart that this only introduces an error of about 1 per cent. into the final net rate of loss of heat. In practice therefore it is sufficient to have observations up to the top of the surface inversion which is practically at 80 m. in the present example.

The rates of loss of heat by various air layers, calculated from the above data in the manner indicated, are given in Table 4.

TABLE 4

RATES OF RADIATIVE HEAT LOSS BY THE LOWEST 80 M. OF THE ATMOSPHERE ON CLEAR JUNE NIGHTS  
(0400 G.M.T.) AT LEAFIELD, OXFORDSHIRE†

Air layer (m.)	0-10	10-20	20-40	40-60	60-80	0-80
Rate of loss of heat (millical. cm. <sup>-2</sup> sec. <sup>-1</sup> )	0.15 (0.12)	0.069	0.073 (0.059)	0.062	0.034 (0.027)	0.39 (0.31)

The values shown in brackets in Table 4 have been evaluated using Elsasser's chart and they are seen to be approximately 20 per cent. less than found by the

\* If in all cases the planimetry is done starting down along the isotherm from the point  $(T_B, u_B)$ ,

where  $T_B$  is the temperature of the bottom of the air layer under consideration,

$u_B$  is the water path from the ground to this level,

a positive result indicates a cooling, negative, a warming.

† Calculated from the observations of Johnson and Heywood(9) using the radiation chart.  
Wind at 0400 G.M.T. force 0 or 1 Beaufort.

present method. The carbon dioxide radiation, not adequately taken into account by Elsasser's chart, accounts almost entirely for the difference, a very small part being due to minor differences in the values of the emissivity of water vapour used in the two charts. This represents an appreciable increase in the accuracy of estimates of the radiative effects in air layers near the ground using the new chart.

This example of the application of the radiation chart shows that radiation plays a large part in cooling the air near the ground during the later parts of clear nights in quiet weather. The 0 to 80 m. air layer would cool at slightly more than 1 °F. per hour due to radiation of heat at the rate of 0.39 millical.  $\text{cm.}^{-2} \text{ sec.}^{-1}$  if no other agencies of heat transfer were operative. Johnson and Heywood (9, Tables 3 and 9) observed rates of cooling of this layer averaging about 0.6 °F./hr. for the period 0200 to 0400 G.M.T. on clear June nights when the average wind was somewhat higher (2.5 m./sec. at 13 m.) than for the occasions dealt with in Table 2 (for which, unfortunately, only the 0400 G.M.T. temperatures are available).

Radiative cooling of the lowest few metres of the air of similar magnitude to the above is also found for a clear calm June night at Porton, Wiltshire, from observations given in some detail by Johnson (11, Fig. 17). The grass minimum temperature on this occasion was 31 °F. and the vapour pressure at 1.2 m. was 7.5 mb. (data kindly supplied by the Meteorological Section, C.D.E.E., Porton). The temperature profile at 0350 G.M.T. (a few minutes before sunrise when the minimum temperatures occurred) was :

Height (m.)	..	..	0	0.6	1.2	2.8	7.1	17.1
Temperature (°F.)	..	..	31	36.9	38.0	39.15	40.2	42.5

from which, using the radiation chart, it is found that the radiative heat loss of the lowest 7.1 m. amounts to some 0.066 mcal.  $\text{cm.}^{-2} \text{ sec.}^{-1}$  which is smaller than the corresponding value of about 0.12 for the Leafeld clear nights, but the difference is no greater than would be expected for different occasions at sites with differing thermal characteristics of the soil.

More complete data, including radiation measurements and observations of the heat lost by the soil, are needed to elucidate satisfactorily the parts played by radiation, conduction, and turbulent heat transfer in cooling the air near the ground under various conditions of wind speed, soil characteristics, etc. It is, however, evident from the above results that radiative cooling of the air must be an important factor in the formation and maintenance of surface temperature inversions and the radiation chart provides an additional tool for the quantitative study of this and allied problems.

#### IV. ADDENDUM

The author is grateful to the referee of this paper for drawing his attention to a chart by Möller(12) which is also designed to include the carbon dioxide radiation.

However, this chart is not readily available nor does it appear to be at all widely known in English speaking countries. Furthermore, the present method

possesses the practical advantage of requiring a single plotting of data and evaluation of areas, whereas in Möller's method both these steps have to be duplicated. The scale of Möller's chart is too small to be suitable for the evaluation of the radiative effects in air layers of the order of 10 m. thickness and the requisite tabulated data for drawing enlarged versions are not available.

#### V. REFERENCES

- (1) SIMPSON, G. C.—*Mem. R. Met. Soc.* **2** (16): 69 (1928).
- (2) SIMPSON, G. C.—*Mem. R. Met. Soc.* **3** (23): 1 (1929).
- (3) BRUNT, D.—*Proc. Roy. Soc. A* **124**: 201 (1929).
- (4) BRUNT, D.—*Proc. Roy. Soc. A* **130**: 98 (1930).
- (5) ELSASSER, W. M.—Harvard Met. Stud. No. 6 (1942).
- (6) BROOKS, F. A.—*Pap. Phys. Ocean. Met. Mass. Inst. Technol.* **8**: 2 (1941).
- (7) ELSASSER, W. M.—*M.W. Rev.* **69**: 1 (1941).
- (8) CALLENDAR, G. S.—*Quart. J. R. Met. Soc.* **66**: 395 (1940).
- (9) JOHNSON, N. K., and HEYWOOD, G. S. P.—*Geophys. Mem.* No. 77, London (1938).
- (10) ROBINSON, G. D.—*Quart. J. R. Met. Soc.* **73**: 127 (1947).
- (11) JOHNSON, N. K.—*Geophys. Mem.* No. 46, London (1929).
- (12) LINKE, F., and MÖLLER, F.—“*Handbuch der Geophysik.*” Vol. 8, Ch. 11. (Berlin, 1943.)



# THE SUPERSTRUCTURE IN THE $\alpha$ PHASE OF SILVER-MAGNESIUM ALLOYS

By L. M. CLAREBROUGH\* and J. F. NICHOLAS\*

[Manuscript received December 15, 1949]

## Summary

X-ray and electrical resistivity studies of a silver-magnesium alloy containing 25 atomic per cent. of magnesium are described. From the results, it is concluded that a superlattice exists at this composition, the order-disorder transformation occurring between 386 and 389 °C.

## I. INTRODUCTION

Andrews and Hume-Rothery(1) suggested that a superlattice should exist at the composition  $\text{Ag}_3\text{Mg}$  in the face-centred cubic  $\alpha$  phase of silver-magnesium alloys, but Letner and Sidhu(2) in their X-ray investigation failed to find support for this prediction.

In the present paper evidence, based on X-ray and electrical resistivity studies, is brought forward to show the existence of a superlattice in an alloy containing 25 atomic per cent. of magnesium.

## II. EXPERIMENTAL

### (a) *Preparation of the Specimens*

Wires of the alloy were prepared as described elsewhere(3). For all annealing treatments the wires were sealed in evacuated Pyrex tubes. The composition of the alloy used was 24.3 atomic per cent. magnesium and the total impurities were less than 0.2 per cent.

### (b) *X-Ray Experiments*

Specimens for powder photographs were prepared by annealing wires of 0.4 mm. diameter for 30 minutes at 550 °C. One specimen was quenched and the other slowly cooled to room temperature at a rate of 30 °C. per hour. After this treatment the wires were etched slightly in nitric acid to remove the surface layers from which some magnesium might have been lost by evaporation. Powder photographs of these specimens, taken with unfiltered copper radiation in a 14 cm. camera, are shown in Plate 1.

Plate 1 (a) shows that the quenched alloy is a disordered face-centred cubic solid solution with a parameter of 4.111 Å. This parameter is larger than that obtained by Letner and Sidhu. Plate 1 (b) shows the diagram obtained from the slowly cooled alloy. Extra lines are clearly visible. Since microscopic examination showed that the alloy was still single phase, in agreement with

\* Division of Tribophysics, C.S.I.R.O., University of Melbourne.

the equilibrium diagram (1), the presence of the extra lines indicates that some ordering process has occurred. The unit cell of the ordered structure, as deduced from Weissenberg photographs (A. McL. Mathieson, personal communication) is a cube having a side length of approximately 16 Å and containing, in all, 256 atoms. The actual arrangement of the atoms in the unit cell has not yet been determined. From powder photographs, the actual side length of this cell was found to be  $4 \times 4.108$  Å. It can be seen from these parameters that the lattice as a whole has suffered a linear contraction of approximately 0.1 per cent. on ordering, which is of the same order of magnitude as that found for  $\text{Cu}_3\text{Au}$  by Owen and Sim(4).

(c) *Effect of Quenching Temperature on Electrical Resistivity*

As one method of determining the transformation temperature, wires were annealed for 30 minutes at 600 °C., slowly cooled at a rate of 30 °C. per hour to various temperatures, and then quenched. Resistances were measured using a

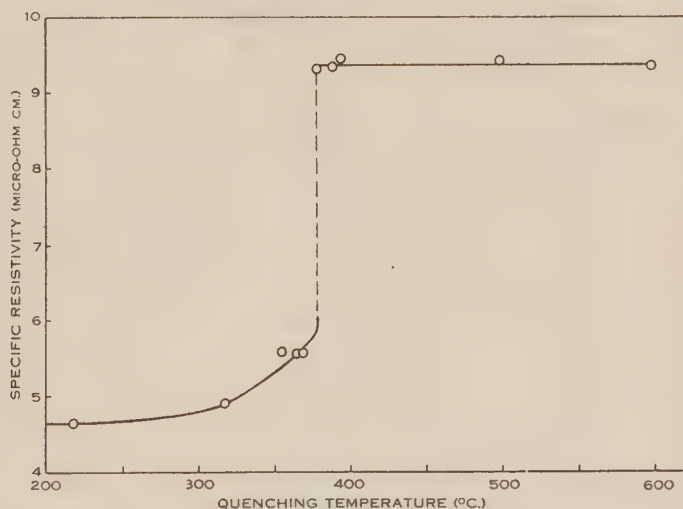


Fig. 1.—Effect of quenching temperature on specific resistivity of  $\text{Ag}_3\text{Mg}$ .

Kelvin double bridge and Figure 1 shows the variation of room-temperature resistivity with quenching temperature. Room temperature varied between 14 and 18 °C. during the course of the experiment. These results indicate that the transformation temperature is approximately 380 °C.

(d) *Variation of Electrical Resistivity with Temperature*

In order to adduce further evidence for the existence of the superstructure and to obtain the transformation temperature more accurately, the resistivity of the alloy was measured at various temperatures.

Owing to the nature of the alloy, spot-welding of current and potential leads to the experimental wire was not possible and silver-soldering was found

to be the most convenient method, although it did cause a 1 per cent. increase in the resistance of the specimen. This error, however, does not affect the determination of the transformation temperature. For a quenched specimen the heating involved in silver-soldering precluded the use of this method and screwed contacts for current and potential leads had to be used. For both methods, a wire 0.97 mm. diameter was bent into a U-form so that it occupied a length of 12 cm. and the current and potential leads were attached. The whole was enclosed in a Pyrex tube of 1.2 cm. diameter with a few turnings of the same alloy to prevent excessive loss of magnesium from the wire and the

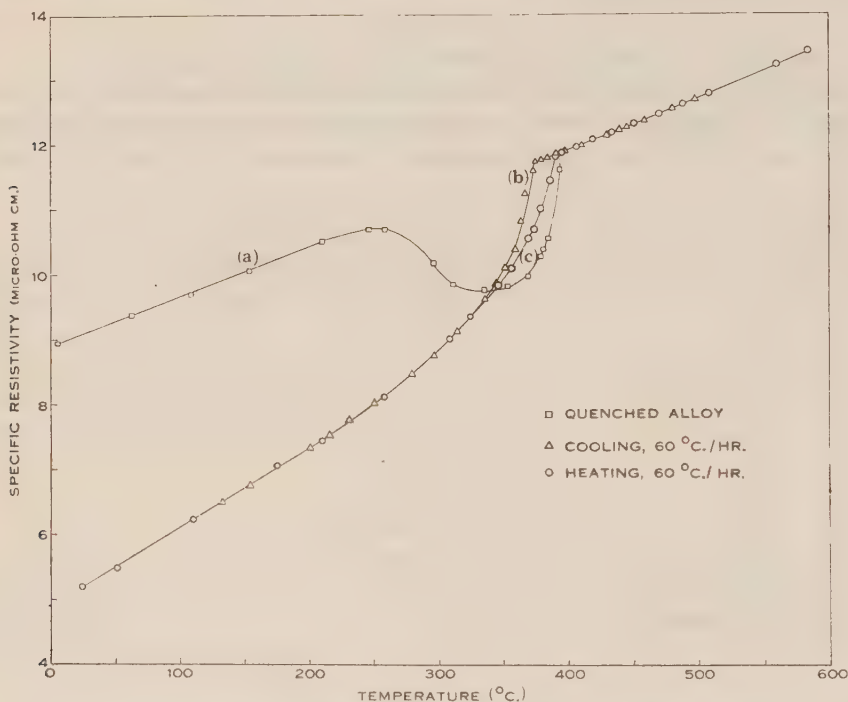


Fig. 2.—Variation of specific resistivity of  $\text{Ag}_3\text{Mg}$  with temperature.

leads were brought out through a rubber stopper. The stopper was made vacuum tight with Apiezon wax and the assembly was evacuated through a side tube and sealed off at  $10^{-4}$  mm. of mercury pressure. The Pyrex tube was inserted in a thick-walled copper tube in the centre of a tube furnace. Temperatures were measured with a chromel-alumel thermocouple, the junction of which was fixed outside the Pyrex tube opposite the centre of the specimen. The dead space in the furnace was packed with asbestos rope and no variation of temperature along the length of the specimen could be detected.

The results are shown in Figure 2. Curve (a) was obtained with a wire quenched from 550 °C. and heated at a rate of 120 °C. per hour. Reordering is indicated by the drop in resistivity, and it can be seen that this commences at about 250 °C. and complete disorder returns at 395 °C. Curves (b) and (c)

were obtained with a second specimen which was carried through several heating and cooling cycles and gave reproducible results for constant heating and cooling rates of 60 °C. per hour. When heating at a rate of 60 °C. per hour, the alloy is completely disordered at 389 °C., and when cooling at the same rate, it starts to reorder at 373 °C.

In the temperature range between 370 and 390 °C. slower heating and cooling rates were then used in order to locate the transformation temperature more accurately. On holding the temperature at 386 °C. for 30 minutes during cooling the resistivity dropped steadily, indicating that the alloy was reordering. Complete disordering could not be obtained below 389 °C., even when the temperature was kept constant for periods of up to one hour in the range 386–389 °C. The large hysteresis shown in Figure 2 is thus mainly due to the rapid heating and cooling rates used, and it appears that the transformation temperature lies in the range 386–389 °C. The lower value found in the quenching experiments above is due to the fairly rapid rate of cooling, viz. 30 °C. per hour, used with this method.

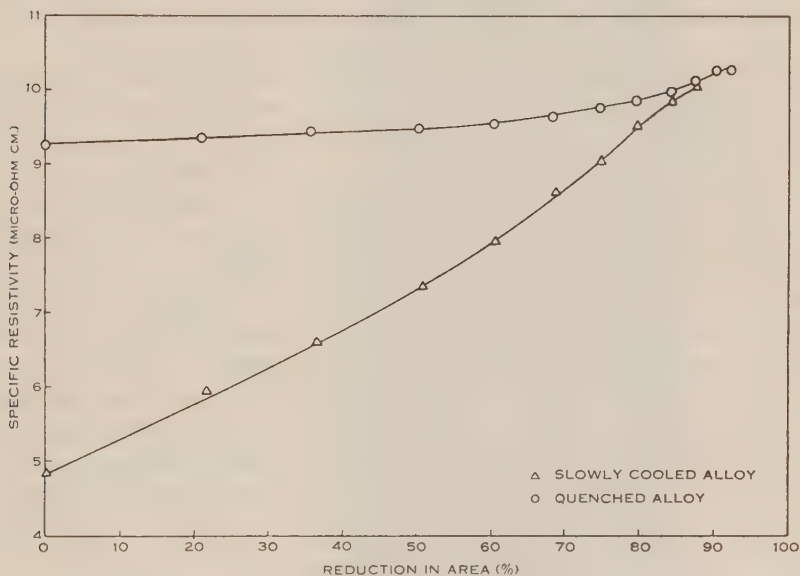


Fig. 3.—Effect of cold work on specific resistivity of quenched and slowly cooled  $\text{Ag}_2\text{Mg}$ .

#### (e) Effect of Cold Work on Electrical Resistivity

To study the effect of cold work on slowly cooled and quenched alloys, two wires 3.66 mm. diameter were annealed at 600 °C. for 16 hours. One wire was quenched from 600 °C. and the other slowly cooled to room temperature at a rate of 30 °C. per hour. Both wires were drawn through carbide dies to approximately 90 per cent. reduction in area and the resistivity measured at various stages. The room temperature throughout the experiment was between 11 and 15 °C. Figure 3 shows the variation in electrical resistivity of the two



wires. It can be seen that the slowly cooled alloy is almost completely disordered after 90 per cent. reduction in area.

The resistivities of the quenched and slowly cooled wires after 90 per cent. reduction in area were measured again when the wires had been kept at room temperature for six months and no change had occurred. This result indicates that, after heavy deformation, order is not restored in  $\text{Ag}_3\text{Mg}$  at room temperature.

### III. DISCUSSION

The occurrence of the  $\text{Ag}_3\text{Mg}$  superlattice is not surprising as conditions are very favourable for it(5). Firstly, the atomic diameters of silver and magnesium are 2.88 and 3.19 Å. respectively, giving a size factor which is small enough to allow a wide range of solid solubility and yet large enough to cause appreciable relief of strain on ordering. Secondly, the electrochemical factor is high, due to the strong electropositive character of magnesium. In the present work, no difficulty was experienced in obtaining the ordered structure and it was probably due to the use of rapid cooling rates, by which the transformation can be suppressed, that Letner and Sidhu(2) were unable to find the ordered state.

In  $\text{Ag}_3\text{Mg}$ , as in  $\text{Cu}_3\text{Au}$ , the transformation can be suppressed by quenching and the resultant disordered state is stable at room temperature. However, on heating quenched (disordered)  $\text{Ag}_3\text{Mg}$  at a rate of 120 °C. per hour it commences to reorder at approximately 250 °C., as compared with approximately 310 °C. for  $\text{Cu}_3\text{Au}$ (6), although the transformation temperatures are almost the same. As the effect of the size factor is similar in each case, this result suggests that the higher electrochemical factor in  $\text{Ag}_3\text{Mg}$  causes the increased tendency to order. This is supported by the fact that much higher deformations are needed to produce comparable disorder in  $\text{Ag}_3\text{Mg}$  than in  $\text{Cu}_3\text{Au}$ . The corresponding deformations are 90 per cent. (cf. Fig. 3) and 60 per cent.(7) respectively.

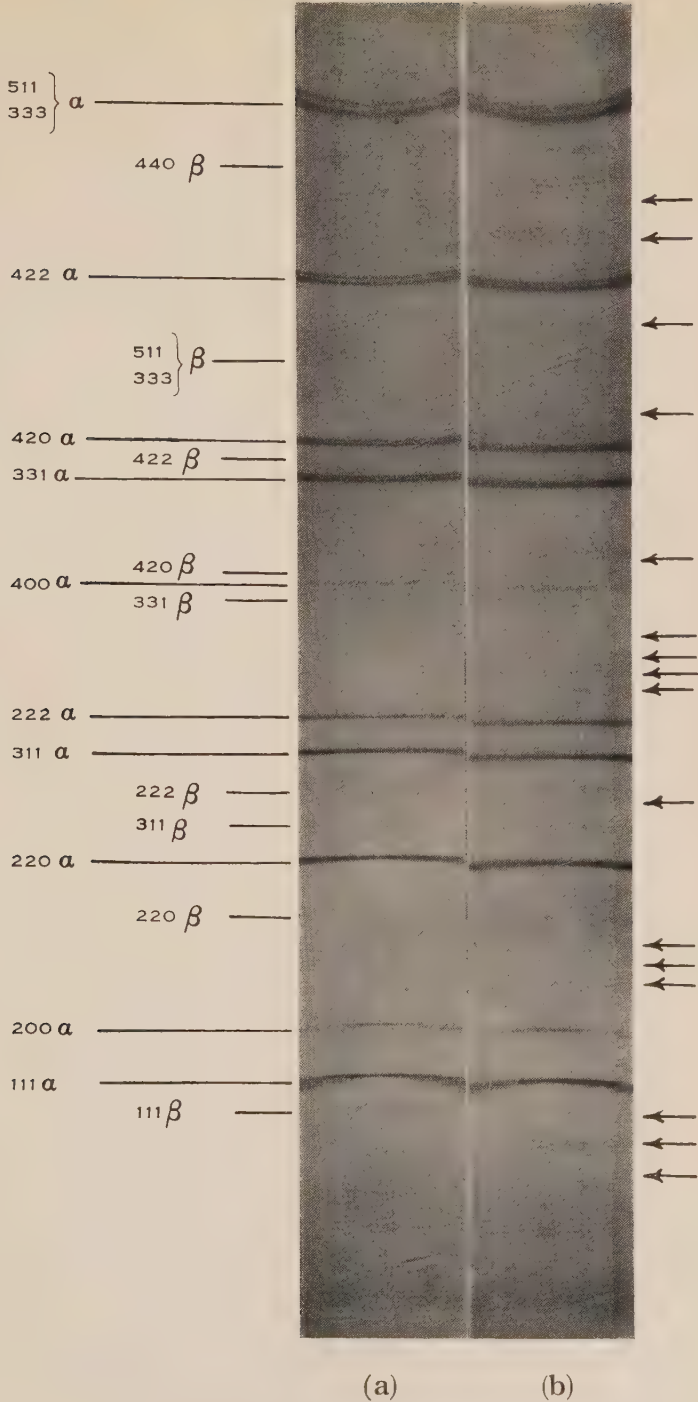
It can be seen from Figure 2 that the transformation temperature given by heating a quenched alloy is slightly higher than that obtained by heating and cooling cycles on an ordered alloy. This behaviour has also been observed for  $\text{Cu}_3\text{Au}$ (6), but is probably only a hysteresis effect due to the heating rates used.

### IV. ACKNOWLEDGMENTS

The authors wish to thank Dr. W. Boas, for his interest in the work and for helpful discussions. They are grateful to Dr. R. I. Garrod, Defence Research Laboratories, Maribyrnong, for taking the powder photographs; to Dr. A. McL. Mathieson, Division of Industrial Chemistry, C.S.I.R.O., for taking Weissenberg photographs from which the ordered structure is being deduced: and to Professor E. J. Hartung for laboratory facilities in the Chemistry Department, University of Melbourne.

### V. REFERENCES

- (1) ANDREWS, K. W., and HUME-ROTHERY, W.—*J. Inst. Met.* **69**: 485 (1943).
- (2) LETNER, H. R., and SIDHU, S. S.—*J. Appl. Phys.* **18**: 833 (1947).



X-ray powder photographs of  $\text{Ag}_3\text{Mg}$  (a) quenched from 550 °C. and (b) slowly cooled. Arrows indicate extra lines.



- (3) CLAREBROUGH, L. M.—*Aust. J. Sci. Res. A* **3**: 72 (1950).
- (4) OWEN, E. A., and SIM, G. M.—*Phil. Mag.* **38**: 342 (1947).
- (5) HUME-ROTHERY, W.—“Atomic Theory for Students of Metallurgy,” p. 111. (Institute of Metals: London, 1946.)
- (6) SYKES, C., and JONES, F. W.—*Proc. Roy. Soc. A* **157**: 213 (1936).
- (7) DAHL, O.—*Z. Metallk.* **28**: 133 (1936).



# THE CATALYTIC DEHYDRATION OF 2,3-BUTANEDIOL TO BUTADIENE

## II. ADSORPTION EQUILIBRIA

By M. E. WINFIELD\*

[*Manuscript received November 16, 1949*]

### *Summary*

Further understanding of the reaction in which 2,3-butanediol is dehydrated over thorium oxide to methyl vinyl carbinol and butadiene has been gained by measuring the adsorption on thoria of water, methyl ethyl ketone and the above-mentioned diol, carbinol, and diene, at temperatures up to 200 °C. In some instances satisfactory isotherms could not be obtained because sufficient chemisorption occurred, followed by polymerization or dehydration reactions, to reduce seriously the area of thoria surface available for adsorption. There was evidence to suggest that water, alone of the vapours concerned, was taken up in greater amount than could be accounted for by adsorption. The bearing of this on the dehydration and its relation to the structure of the catalyst is briefly discussed. A suggestion in a previous paper(1) that water is the product which retards the catalytic dehydration is confirmed.

At temperatures little above 50 °C. methyl vinyl carbinol is dehydrated by thoria to butadiene, while methyl ethyl ketone at room temperature undergoes self-condensation on thoria to yield a methyl heptenone. Also at quite low temperatures dimerization of butadiene brings about appreciable reduction of available surface. Some loss of activity on this account must therefore be expected when thoria is used to dehydrate 2,3-butanediol, even though butadiene is adsorbed only weakly.

## I. INTRODUCTION

In a previous paper(1) it was shown that 2,3-butanediol can be dehydrated in the presence of thoria to methyl vinyl carbinol, and butadiene. A preliminary investigation of the kinetics of the reaction was described. Later a detailed study of the mechanism of the reaction was commenced, aimed at extending the understanding of gas-phase catalytic dehydration. Its mechanism is poorly understood compared with that of hydrogenation, perhaps because the catalyst and reactants can be simpler in the latter case. The findings of the investigation were expected to lead to improved yields of butadiene.

It has been suggested(1) that water is the reaction product whose slow removal from the adsorbent retards the overall catalytic process. No technique was available for determining desorption rates and thus confirming this directly. Confirmation has been provided indirectly by adsorption equilibria measurements, which were undertaken for the following reasons:

Some investigators(2) assume that in a catalysis that is poisoned by a product, adsorption equilibrium between product and surface is actually attained, under "steady state" conditions. The fraction of surface covered by poison

\* Division of Industrial Chemistry, C.S.I.R.O., Melbourne.

during the catalysis is then equal to the equilibrium amount, which can be obtained from the adsorption isotherms. In any case, the heats of adsorption can be calculated from the isotherms, and it is the heat of adsorption that largely governs the rate of desorption, the activation energy for adsorption being considered low(3). Adsorption equilibrium measurements, particularly if carried out at the highest practicable temperature, were expected to provide information regarding the relative ease of chemisorption of the compounds, though it was realized that most of the observed adsorption would be van der Waals'. A matter of some interest was to determine the relative stability of MVC, butadiene, and MEK when adsorbed, because it was found in previous flow experiments that carbon appeared on the catalyst during use.

To complement the adsorption measurements it was necessary to investigate the structure of the catalyst, in search of an explanation for the specificity of thoria in catalyzing the conversion to MVC and the diene, and an explanation of the difficulty in desorbing the retarding product.

## II. EXPERIMENTAL PROCEDURE

Adsorption equilibria for butadiene were measured using an apparatus for the determination of surface area and described by Brunauer, Emmett, and Teller(4). The butadiene was prepared from *cyclohexane*(5). It was converted to the tetrabromide, which after three crystallizations from ethanol was heated with zinc dust and ethanol to regenerate the diene. A column of eight theoretical plates was used for distillation of the product, and the fraction which boiled steadily at  $-4.5^{\circ}\text{C}$ . at 760 mm. was retained for the adsorption measurements.

In the experiments with adsorbates that are liquids at room temperature the adsorption cell consisted of a flat-bottomed Pyrex glass vessel, of volume 230 cc., in which 3.00 g. of adsorbent was spread in a thin layer. Attached to the cell was a spoon gauge, and in addition a horizontal side-arm in which was placed an ampoule containing the adsorbate. The neck of the ampoule passed through a glass collar, and was fractured by turning the collar which for this purpose was connected vertically through a ground joint to an external lever. A projection-viewing device greatly facilitated observation of the needle of the spoon gauge. A "Micromax" recorder-controller with a multi-junction thermocouple regulated the temperature of an oil-bath in which the cell was deeply immersed.

With this type of adsorption apparatus, each run results in the determination not of a number of points on a single isotherm, but one point on the isotherm for each of a succession of increasing temperatures. The complete isotherm for each temperature is attained after a series of runs, each with a different weight of adsorbate. For this reason the isotherms show clearly the degree of experimental error. Whether each of the points lies on a smooth curve depends upon the temperature of the degassing procedure being identical for each run, and upon the adsorbent suffering no change in surface area and activity from run to run.

Adsorbate liquids were purified by drying, distilling through a column of 15 theoretical plates, and finally by three distillations at a pressure of less than

1 micron. The final distillate collected in an ampoule which was sealed off under high vacuum, and which, before and after filling, was weighed along with the small ground joint through which connection was made to the vacuum system. All of the thorium oxide was taken from a batch designated C30, which was prepared according to the method described previously(1) for C24 except that the oxalate was ignited at 400 °C. for six hours. Before each run the adsorbent was degassed at 270 °C. for four to six hours at a pressure of less than 0.1 micron.

### III. EXPERIMENTAL RESULTS

When calculating the amount of gas adsorbed it was assumed that the volume of a mole of the gas under standard conditions was 22.4 l. except for butadiene, for which the value 21.81 l.(6) was employed.

#### (a) Water Adsorption

Adsorption isotherms for temperatures from 50 to 200 °C. are shown in Figure 1. Dotted curves represent the best lines through the experimental

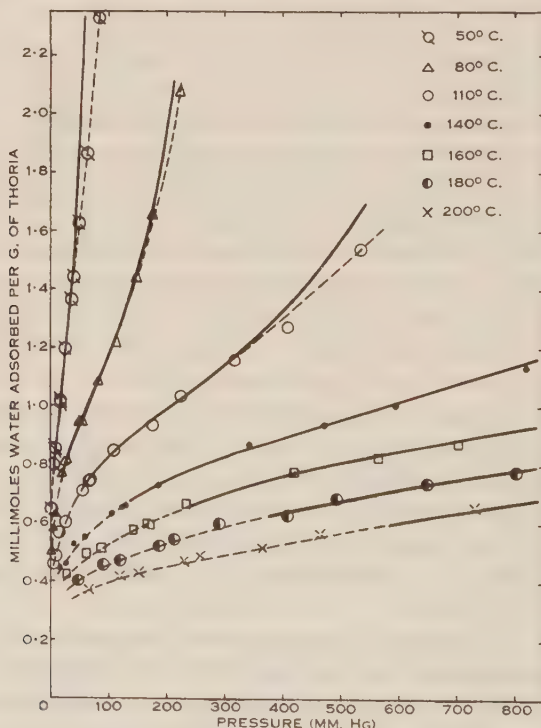


Fig. 1.—Sorption isotherms for water on thorium oxide.

points. The upper half of the experimental curve for 50 °C. can be described by the equation

$$\log_{10} \frac{p}{p_0} = 0.162 - \frac{1.139}{m^2},$$

which is in the form given by Harkins and Jura(7) for a condensed film ( $p$  is the pressure in mm. Hg,  $p_0$  the saturation pressure at a given temperature, and

$m$  is the number of millimoles of vapour adsorbed per gram of thoria). From this equation the surface area of the thoria was calculated to be 95 square metres per gram, a value which later is suggested to be in error. Harkins and Jura's constant  $k$  was estimated for the purpose of the above calculation to be 3.99 for water at 50 °C.

The full lines in Figure 1 are drawn according to an equation of the form given by Brunauer, Emmett, and Teller(8)

$$\frac{p}{m(p_0 - p)} = \frac{1}{0.88d^{2/3}e^{168/RT^{1/2}}} + \frac{e^{168/RT^{1/2}} - 1}{0.88d^{2/3}e^{168/RT^{1/2}}} \cdot \frac{p}{p_0}, \dots (1)$$

where  $d$  is the density of water at the temperature  $T$  °K. In the pressure range for which the B.E.T. equation generally holds, from  $p/p_0 = 0.05$  to  $0.35$ , the equation represents the adsorption of water on thoria C30 with a probable error of about 2 per cent. ( $E_1 - E_l$ ) derived from the equation was 3300 cal. per mole at 180 °C.,  $E_l$  being the heat of liquefaction of the adsorbate and  $E_1$  the average heat of adsorption in the first layer of molecules adsorbed.

Isosteric heats of adsorption,  $E_i$ , were calculated from the experimental isotherms and plotted as a fraction of  $\theta_e$ , the fraction of the surface apparently covered by adsorbate. Towards low values of  $\theta_e$  the curve was approximately linear. Though this could not be extrapolated to  $\theta_e = 0$  with any exactitude, it may be taken that the gross heat of sorption for the first molecules of water that are sorbed is in the vicinity of 24,000 cal. per mole. Van der Waals' forces cannot account for such a large value, which suggests that when  $\theta_e$  is small sorption involves ionization of the water molecules.

TABLE I

Adsorbate	Molecular Area from Emmett Equation (sq. Å)	Molecular Area from Adsorption Isotherms (sq. Å)	Temperature (°C.)
Water .. ..	10.6	4.6	50
Diol .. ..	31	33	160
Butadiene .. ..	30	40	50
MVC .. ..	30	236	50

The surface area available for water adsorption can be calculated from equation (1) provided a value can be obtained for the area occupied by an adsorbed water molecule. It is assumed for this purpose that most of the molecules adsorbed in a complete unimolecular layer are not ionized, which would appear to be the case because ( $E_1 - E_l$ ) is only about 3000 cal. per mole. The known atomic radii of hydrogen and oxygen lead to a value in the neighbourhood of 11 sq. Å for the area of a molecule. When calculated from Emmett's equation(9) the figure 10.6 sq. Å given in Table 1 was obtained. Taking the



latter value, the surface area of the adsorbent was calculated to be 56 sq. m./g. whereas the area was only 24.3 sq. m./g. when determined by the B.E.T. method using nitrogen as adsorbate at liquid oxygen temperature.\*

### (b) Butadiene Adsorption

Some of the isotherms are shown in Figure 2. Those for 50 and 80 °C. were obtained by plotting the experimental values of  $p$  and  $m$  up to the point where they began to deviate from the anticipated curve, due to butadiene polymerization, and then extrapolating to higher pressures by extrapolating B.E.T.(8) straight line plots of the experimental values of the lower pressures. The  $p/p_0$  limits of applicability of the B.E.T. equation were 0.07 to 0.3. Values of  $p_0$  were calculated from the equation of state for butadiene(10)

$$\log_{10} p_0 = \frac{-973.6}{t \text{ } ^\circ\text{C.} + 243.2} + 6.96128.$$

It will be noticed in Figure 2 that the amounts adsorbed at 110 °C. are lower than might be anticipated from a consideration of the isotherms at lower

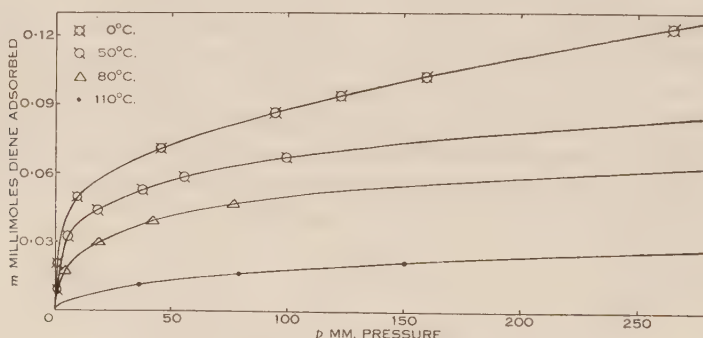


Fig. 2.—Butadiene adsorption isotherms, corrected, except at 0 °C., for dimerization.

temperatures. Though the experimentally determined pressures were corrected for the appreciable dimerization which occurs at 110 °C. (by measuring the rate of polymerization at each pressure) it was not possible to take into account the reduction in surface area of adsorbent brought about by deposition of dimer. At 0 °C. dimerization was negligible and the isotherm could be determined accurately. It is only for this temperature that the points shown in the figure are the uncorrected experimental points. A reliable value of  $E_i$  cannot be derived from the temperature dependence of the adsorption. At 0 °C. ( $E_1 - E_i$ ) was 2000 cal. per mole, which is appreciably less than for water or diol.

### (c) 2,3-Butanediol Adsorption

Earlier work(1) indicated that the catalytic dehydration of the diol in the presence of thoria would be inappreciable at 250 °C. Nevertheless, it was found in the diol adsorption runs at temperatures up to 200 °C. that the available

\* Carried out by Mr. H. Hergt, Division of Industrial Chemistry.

surface decreased markedly from one run to the next, until three or four runs had been carried out with the same sample of adsorbent. This will be evident from the isotherms of Figure 3, where the full lines refer to the first run with fresh adsorbent, and the broken lines to succeeding runs. It was interpreted to mean that even at temperatures below 200 °C. chemisorption of diol took place to some extent, leading to reactions which caused part of the surface of the thoria to be covered with compounds that are removed only with difficulty.

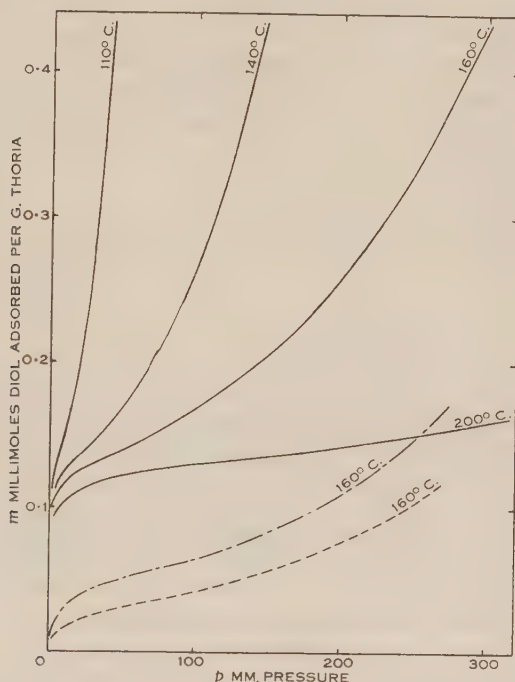


Fig. 3.—Approximate 2,3-butanediol adsorption isotherms, obtained by calculation, showing decrease in surface area of adsorbent in successive runs.

As a result few of the 18 diol runs were reliable. Isotherms which were approximately correct, however, were constructed by first plotting the results of run G41 (a run with fresh adsorbent) according to the B.E.T. equation. The curves thus obtained for each of the temperatures from 110 to 200 °C. were so close together that they appeared to fall on one line (straight between  $p/p_0 = 0.05$  to  $0.35$ ) and this line was used in calculating the whole range of  $m$  and  $p$  values (full lines in Fig. 3) at each of the temperatures 110, 140, 160, and 200 °C. Reconstructed 160 °C. isotherms based on run G42, and on runs G43 and G44 (the results of these two fell on the same B.E.T. curve) appear as broken lines in the figure. It is appreciated that extrapolation from B.E.T. plots of the kind resorted to above is a questionable procedure. The results of the diol adsorption measurements have been presented graphically for convenient, if rough, comparison with results for other adsorbates.

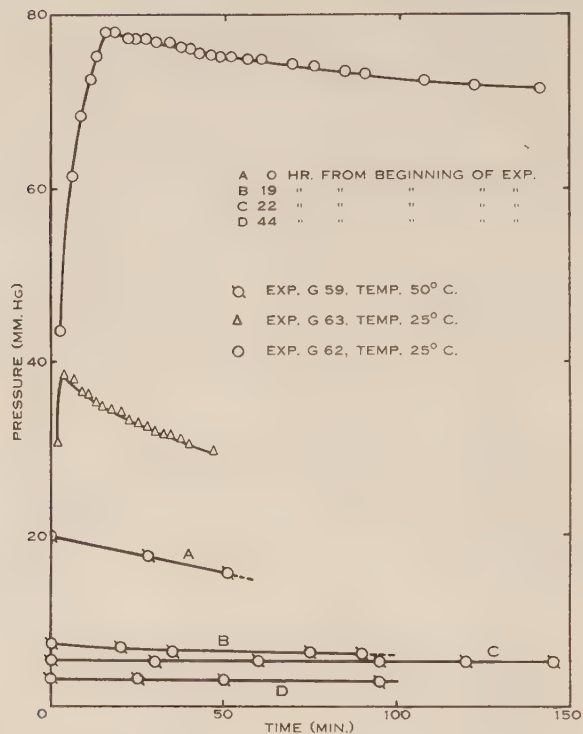


Fig. 4.—Apparent rate of adsorption of methyl ethyl ketone on thoria.

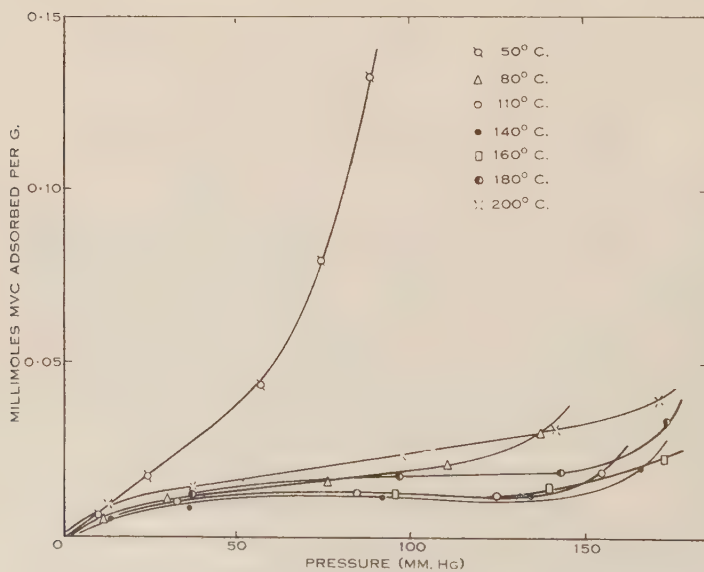


Fig. 5.—Methyl vinyl carbinol isotherms illustrating chemisorption at the higher temperatures.

(d) *Methyl Ethyl Ketone Adsorption*

Several attempts to measure adsorption isotherms for MEK on thoria were unsuccessful because an equilibrium could not be attained even at room temperature. After releasing MEK in the adsorption cell in the presence of thoria the pressure continued to fall for many hours, as shown in Figure 4. The initial rise in pressure was due to the relatively slow boiling of MEK out of the ampoule at low temperatures. An adsorption rate determination at 50 °C. and a few mm. pressure, covering the first 20 seconds after admission of adsorbate, resulted in a curve similar in shape to those in Figure 4. From these preliminary experiments it was concluded that MEK is chemisorbed on thoria at low temperatures, and that following adsorption a chemical reaction occurs, presumably a polycondensation, resulting in a product of low vapour pressure, little of the water that is formed being desorbed. The nature of the reaction is referred to in a later section.

If chemisorption of MEK is not conceded, it must be assumed that the condensation is a homogeneous reaction and that thoria influences the equilibrium by adsorbing the water that is formed.

(e) *Methyl Vinyl Carbinol Adsorption*

Figure 5 shows the amounts of MVC apparently adsorbed on thoria at temperatures from 50 to 200 °C. when the adsorbent was degassed at 270 °C. before each run. When degassing was carried out at 220 °C. a different set of isotherms resulted in which the amounts adsorbed were approximately one-third of those shown, except at 50 °C. where they were only a little less. A third series of curves, in which the adsorption was consistently greater than that shown in the figure, was obtained when the degassing temperature was 290 °C. It was not practicable to degass at higher temperatures with the type of adsorption cell employed in these investigations.

It has been suggested\* that the 50 °C. curve in Figure 5 may be a type III isotherm(11), so that the value for the apparent area of an adsorbed MVC molecule calculated by the B.E.T. method and given in Table 1 may be in error. Yet it does show clearly enough that in the experiments to which Figure 5 refers only a small proportion of the surface was available for MVC adsorption. Later it is shown that the carbinol will yield butadiene in the presence of thoria at 100 °C. Thus butadiene dimer or some product of its thermal decomposition was probably responsible for the surface poisoning. Increasing the temperatures of degassing presumably reduced the fraction of surface that was masked. If the degassing temperature was kept constant, the isotherms at 50, 80, and 200 °C. were made up of experimental points which lay on a smooth curve, even though each point was obtained in a different run. The uncertainty that is apparent in the measurements at 110, 140, 160, and 180 °C. is not understood, nor the minima in the curves, but it is believed to be a consequence of diene formation.

\* Dr. P. H. Emmett and Dr. T. W. de Witt of the Mellon Institute of Industrial Research kindly criticized the manuscript.



At temperatures above 140 °C. the amounts apparently adsorbed increased with increasing temperature. Irrespective of whether this was due to increased chemisorption while van der Waals' adsorption dwindled, or to increased dimerization of butadiene, it can be concluded that MVC was chemisorbed at those temperatures.

(f) *Reaction of Methyl Ethyl Ketone Chemisorbed on Thoria*

MEK was refluxed at atmospheric pressure for several hours with thorium oxide, while the extent of reaction was measured qualitatively by determining the refractive index of the liquid at frequent intervals. The water that was formed progressively slowed the reaction by poisoning the catalyst. It was therefore attempted in later experiments to remove some of the water by continuous azeotropic distillation with benzene. At the conclusion, the benzene and unreacted MEK were removed by distillation at one atmosphere. Fractional distillation of the residue at 60 mm. yielded as principal fraction a colourless liquid with a pleasant fruity odour. At 761 mm. pressure it boiled at 165.9 °C. The refractive index was  $n_D^{21.8}$  1.4479. Neither a crystalline phenylhydrazone nor semicarbazone could be obtained. However, infra-red spectroscopy\* showed that the compound was an unsaturated ketone. When compared with MEK there was a new band corresponding to a carbon-carbon double bond, with no trace of the saturated ketone. Examination in the mass spectrometer† established the molecular weight as 126 and the weight of the principal fragments as 97, 69, 57, and 29.

The fragment of weight 57 suggested that the molecule contained the group  $C_2H_5\overset{|}{C}=O$  and the presence of this rather than of  $CH_3\overset{|}{C}=O$  was supported by a negative iodoform test. The liquid was taken to be either 3-methyl-3-heptenone-5 or 3-methyl-2-heptenone-5 or a mixture of both(12).

(g) *Reaction of Methyl Vinyl Carbinol Chemisorbed on Thoria*

When MVC was refluxed at atmospheric pressure with thorium oxide there was no change in refractive index, but in each of several experiments butadiene was evolved shortly after boiling commenced. When formation of the diene had ceased the liquid product was separated by distillation at one atmosphere pressure into three fractions. The first, collected at 90–102 °C., was largely MVC. A small second fraction then came over at 102–110 °C. The temperature was of little significance since distillation was on a small scale, and proceeded very slowly. A residue of butadiene dimer remained in the still. The second fraction was not large enough to allow purification and identification with certainty of several solid derivatives that were prepared. However, the principal constituent was shown to be an alcohol, so that there are only two alternatives for its identity that need be considered, 2-butene-1-ol and 3-butene-1-ol. The former is favoured because it is known to be formed readily by isomerization of MVC with dilute acid.

\* Carried out by Mr. A. Walsh, Division of Industrial Chemistry.

† Carried out by Mr. G. Hercus, Division of Industrial Chemistry.

## IV. DISCUSSION AND CONCLUSIONS

*(a) General*

Most of the adsorption that was measured was van der Waals', indicated by low average heats of adsorption,  $E_1$ , at all temperatures up to 200 °C. Where there was evidence of chemical reactions occurring at the surface when diol, MEK, or MVC was adsorbed, it was considered that a chemisorption process preceded reaction. At the lowest temperatures at which reaction may have taken place little evidence of its occurrence was expected, because practically no desorption of the water formed could occur, water being sorbed in considerably greater amounts than other adsorbates. Several practical difficulties have prevented determination of low-pressure isotherms. No conclusion regarding chemisorption of butadiene was reached, because the dimerization that was observed is known to occur as a homogeneous reaction, although more slowly. It is controversial whether van der Waals' adsorption could account for the acceleration, perhaps by increasing the entropy of activation for the dimerization reaction.

The B.E.T. equation(8) for multilayer adsorption proved to be the most useful relation for expressing the results of adsorption on thoria as straight line plots. For lower values of  $p/p_0$  than 0.05 Fowler's(13) statistical equation (No. 9, p. 147) could be applied with advantage. The equation of Harkins and Jura(7) for a condensed film gave straight line plots only at exceptionally high relative pressures when applied to the sorption isotherms for water, perhaps because more than half of the water taken up was by a process other than adsorption. The apparent surface area determined from the Harkins and Jura plots is therefore of doubtful significance.

*(b) Sorption of Water and Its Relation to the Catalytic Mechanism*

Water is sorbed strongly by thorium oxide. When  $\theta_e$  is very small the heat of sorption is approximately 24,000 cal. per mole. There are more than twice as many sites for sorption of water as of other molecules if the comparison takes account of the area of each molecular species (Table 1). Thoria is known to retain some water tenaciously. The last traces can be removed only by heating above 800 °C.(14). The high sorption heat, higher than for water on most other adsorbents, may be accounted for by water held within the thorium lattice. X-ray examination\* of the latter showed the presence of a "superlattice" of double the normal thorium oxide dimensions and corresponding to the lattice of the thorium oxalate, from which the oxide had been produced by thermal decomposition at the lowest practicable temperature.† The uptake

\* By courtesy of Dr. A. L. G. Rees, Division of Industrial Chemistry.

† Temperatures up to 350 °C. were employed when preparing samples for X-ray analysis. It was shown by measuring the electrical conductivity of a pellet of thorium oxalate during decomposition that 340 °C. is adequate. Hahn and Senftner(14) claimed that a much higher temperature was required, but it is believed that their result was due to further removal of water from the lattice and not to oxalate decomposition. The change which they noted at 290–420 °C. probably corresponds to the latter process.

of water by the thoria used in these experiments may resemble the process described by Barrer(15) for zeolites. In the early experiments on catalytic activity, it was found that the thoria declined in activity slowly if used at 500 °C. and fairly rapidly, that is after only a few runs, at 550 °C. Later it was shown that heating the thoria in air at 500 °C. increased the sharpness of the X-ray powder patterns, although thoria cannot be expected to sinter at temperatures much below 1400 °C.(16). It was therefore concluded that the "porous" lattice of the active thoria was stabilized by the presence of water in the "holes" and that removal of some of the water at higher temperatures led to partial collapse to the normal thoria dimensions. Barrer's analogous conclusions with regard to the zeolites strengthen the supposition. Also, the case is not far different from that of metal films whose laminar structure is stabilized by the presence of a gas(17, 18) except that the spaces between the laminae appear to be larger.

An X-ray study of the mode of thermal decomposition of thorium oxalate is planned, similar to those of silver oxalate by Macdonald(19) and by Griffith(20), but has been postponed because of difficulty in growing large enough crystals of the oxalate. Meanwhile, certain of their conclusions are adopted although the decomposition which they studied was chemically different. It is supposed that decomposition commences from a surface site and in particular from a site of high energy, that is, a defect. Once started the decomposition proceeds by diffusion of the resulting "holes" along the particular crystal plane through which diffusion is easiest. When the decomposition plane cuts a surface, new defects are formed from which new planes of decomposition may commence. Thus the thoria crystal finally obtained is envisaged as containing a number of channels along which, or over the surface of which, an ion such as  $\text{OH}^-$  is expected to diffuse readily. Recent sorption rate measurements to be published later have shown that the process is indeed rapid.

Macdonald noted inhibition of the decomposition by nitrate or sulphate, suggesting that these ions influence the points from which decomposition begins or the paths along which it proceeds. This is in accord with the finding(1) that a trace of nitrate or sulphate is sufficient so to affect the decomposition of thorium oxalate that the product is of greatly inferior activity in the dehydration to MVC and butadiene.

There seems no point in attempting yet to correlate the interatomic dimensions of 2,3-butanediol with those of the catalyst, now that it is realized that the latter cannot be assumed to have the normal thoria structure. It is believed that the specific solution of water in the catalyst will prove to be important in any explanation of the special efficiency of thorium oxide in the dehydration reaction. The activation energy for the diffusion of water into the solid (which is being measured at present) can be taken to be considerably less than the activation energy of the surface reaction, so that water removed from diol adsorbed on the surface may be withdrawn from the adsorption site rapidly by inward diffusion. Thus the equilibrium between adsorbed reactant and adsorbed product is displaced in favour of the latter. It is suggested that the



conceptions regarding the "steady state" concentration of hydrogen in palladium during catalytic hydrogenation, presented by Wagner and Haufler(21), can be extended to catalytic dehydration in certain cases.

From the water adsorption isotherms were obtained, by extrapolation, approximate values for  $\theta_e$ , the fraction of the total number of adsorption sites that is covered by water, at adsorption equilibrium at a given pressure and at the temperatures employed(1) in obtaining the experimental results dealing with the influence of pressure and temperature on the catalytic dehydration of 2,3-butanediol. It is then possible to calculate approximately the value of  $\theta$ , the fraction of the "active" surface that is covered with molecules of all species during catalysis under "steady state" conditions.

The conversions  $C$  and  $C'$  are compared at two pressures  $P$  and  $P'$  while temperature and volumetric flow-rate are kept constant.  $v$  is the overall reaction velocity when the total pressure is  $P$ ;  $p$  is the partial pressure of the retarding product, water vapour;  $\theta_0$  is the fraction of the "active" surface that is covered with reactant, in the "steady state". For simplicity, rates of adsorption are taken to be a linear function of  $(1-\theta)$  and of pressure, rates of desorption a linear function of the surface covered by the particular species. It is assumed that the presence of products other than water on the surface can be neglected, that both reactant and product are held by either one point or two point adsorption, and that both are adsorbed on the same type of sites.  $\theta_0$  is taken to be small compared with  $\theta$ , so that the fraction covered by water is approximately  $\theta$ .

While the catalytic reaction proceeds in the "steady state" it cannot be assumed that the retarding product is in adsorption equilibrium with the surface unless  $\theta_e$ /(fraction of total surface that is "active") is practically equal to  $\theta$ . This is believed not to be the case for the diol dehydration at reduced pressure. For the present, the conventional assumption, that the reactant does attain adsorption equilibrium, is retained. Thus  $\theta_0$  is proportional to the free surface and to the partial pressure of the reactant, which is  $(P-2p)$ ; that is

$$\theta_0 = k_1(1-\theta)(P-2p),$$

therefore

$$v = \text{rate of surface reaction} = k_2\theta_0 = k_3(1-\theta)(P-2p),$$

and

$$\begin{aligned} v &= \text{rate of water desorption} - \text{rate of water adsorption} \\ &= k_4\theta - k_5(1-\theta)p. \end{aligned}$$

If one half of the total surface is "active", then the fraction of the "active" surface that is covered by adsorbed water at equilibrium, in the absence of other species of gas, is at most  $2\theta_e$ ,

therefore

$$2\theta_e k_4 = k_5(1-2\theta_e)p,$$

that is

$$v = k_4\theta - \frac{2k_4\theta_e(1-\theta)}{(1-2\theta_e)}.$$



Thus if  $v'$  is the velocity at pressure  $P'$ ,

$$\frac{v'}{v} = \frac{(1-\theta')(P'-2p')}{(1-\theta)(P-2p)} = \frac{\theta'-2\theta_e'(1-\theta')/(1-2\theta_e')}{\theta-2\theta_e(1-\theta)/(1-2\theta_e)}$$

Also

$$\frac{v'}{v} = \frac{P'C'}{PC}$$

Suitable numerical values are available from the previous paper. For example, at 350 °C.,

when  $P=16$  mm.,  $C=33.7\%$ , thus  $p=5.3$  mm.,

and when  $P'=32$  mm.,  $C'=26.6\%$ , thus  $p'=8.5$  mm.

Since  $\theta_e$  is at most 0.002 when  $p=5.3$ ,  $\theta$  can be calculated to be 0.41, and  $\theta'$  to be 0.66. Alternatively, if only 1% of the total surface is "active",  $\theta=0.47$  while  $\theta'=0.69$ .

If the reactant is not in adsorption equilibrium the same values of  $\theta$  are obtained, provided the rate of desorption of reactant is taken to be negligible.

No great accuracy can be claimed for the relation between  $\theta$  and  $P$  that is illustrated in these calculations, but it is of interest for two reasons. The values of  $\theta$  are in accord with the qualitative experimental finding that reducing the pressure does not lead to appreciable improvement in yield except below about 250 mm. and that there is little advantage in reducing the pressure below 8 mm. Also,  $\theta$  must be known at least approximately before an attempt can be made to understand the significance of  $E_{\text{exp}}$ , the experimental activation energy for the overall reaction.

Beeck(22) has drawn attention to the great importance that may attach to the meaning of the entropy and energy of activation in the extraordinary case of a hydrogenation for which the observed activation energy is practically constant for a series of metal catalysts. An authoritative discussion of the meaning of his  $E_{\text{exp}}$  would have been most helpful. Presumably it was made up of three separate energies, because the reaction was retarded by an undesirable chemisorption of one reactant on the metal.

For a reaction which is retarded by slow removal of a product from the surface, several authors (e.g. 2, 3) have suggested the simple equation

$$E_{\text{exp}} = E_0 - E + E_p \quad \dots\dots\dots (2)$$

Using Glasstone, Laidler, and Eyring's(3) nomenclature,  $E_0$  is the activation energy for the reaction on a clean surface,  $E$  is the heat of adsorption of reactant and  $E_p$  the heat of adsorption of product. According to the assumptions made by Hinshelwood(2) the equation applies only when both  $\theta$  and  $\theta_e$ /(fraction of total surface that is active) are close to unity, and therefore is very limited in its application. The second of these conditions has escaped the notice of some investigators and has led to criticism(23) of the equation. Under the experimental conditions in which the activation energy was determined for the diol dehydration,  $\theta_e$  was at most 0.005 and  $\theta$  approximately 0.5. An attempt was therefore made to derive the meaning of  $E_{\text{exp}}$  by assuming with Hinshelwood

that in the "steady state" the reactant reaches adsorption equilibrium, but making no such assumption with regard to the retarding product.

In order to determine the significance of  $E_{\text{exp}}$  it is convenient to examine the temperature dependence of the overall reaction by varying the flow-rate so that  $P$  and  $p$  are independent of  $T$ .

Then

$$v = k_6 \times k_{\text{exp}},$$

where  $k_{\text{exp}}$  is the observed rate constant and  $k_6$  is a constant that is independent of  $T$ . The temperature dependence of  $k_{\text{exp}}$  is then given by the following three equations:

$$k_{\text{exp}} = k_7 e^{-E_{\text{exp}}/RT} \text{ (overall reaction), } \dots\dots\dots (3)$$

$$k_{\text{exp}} = k_8 \theta_0 e^{-E_0/RT} \text{ (surface reaction), } \dots\dots\dots (4)$$

$$k_{\text{exp}} = k_9 \theta e^{-E_p/RT} - k_{10}(1-\theta) \text{ (desorption of product), } \dots\dots (5)$$

the activation energy for adsorption of water being considered negligibly small.

Now that pressure is a constant,  $\theta_0$  will be proportional to the free surface and to the exponential of the heat of adsorption of reactant, if the latter attains equilibrium as is commonly supposed;

that is

$$\theta_0 = k_{11}(1-\theta)e^{+E/RT},$$

and equation (4) can be written

$$k_{\text{exp}} = k_{12}(1-\theta)e^{(E-E_0)/RT}. \dots\dots\dots (4a)$$

When  $\theta_e$ /(fraction of surface that is "active") is close to  $\theta$ , and  $\theta$  close to unity, equations (3), (4a), and (5) lead to equation (2). In the present case, however, the extent of poisoning of the surface, when measurements are taken of the adsorption of 2,3-butanediol or methyl vinyl carbinol, suggests that the fraction of the adsorbent surface that is catalytically "active" is not small. Therefore, since  $\theta_e$  is at most 1 per cent. of  $\theta$ , the rate of adsorption of water is small compared with the rate of its desorption, and equation (5) becomes

$$k_{\text{exp}} = k_9 \theta e^{-E_p/RT}. \dots\dots\dots (5a)$$

With this approximation equations (3), (4a), and (5a) can be solved to give a simple expression for  $E_{\text{exp}}$ . It will be seen that when  $\theta$  is neither small nor large the significance of  $E_{\text{exp}}$  in terms of the other energies cannot be determined without knowledge of a particular value of  $\theta$  and the relevant values of  $E$  and  $E_p$ .

From equations (5a) and (3)

$$\theta = \frac{k_7}{k_9} e^{(E_p - E_{\text{exp}})/RT},$$

while from (4a)

$$\begin{aligned} \frac{d}{dT} \ln k_{\text{exp}} &= \frac{E_0 - E}{RT^2} + \frac{d}{dT} \ln (1-\theta), \\ &= \frac{E_0 - E}{RT^2} + \frac{\theta(E_p - E_{\text{exp}})}{(1-\theta)RT^2}. \end{aligned}$$

But

$$\frac{d}{dT} \ln k_{\text{exp}} = \frac{E_{\text{exp}}}{RT^2}.$$

Therefore

$$E_{\text{exp}} = (1-\theta)(E_0 - E) + \theta E_p. \dots\dots\dots (6)$$

It should be emphasized that equation (6) cannot hold when adsorption equilibrium for the retarding product is close to being attained.

When the reactant also is far from reaching adsorption equilibrium, a simple expression for  $E_{\text{exp}}$  cannot be obtained except when the rate of desorption of reactant is negligible. This may prove to be a common case because if the adsorption of reactant is relatively slow there is a good likelihood of its being an irreversible chemisorption. Equation (4a) is then replaced by

$$k_{\text{exp}} = k_{13}(1 - \theta)e^{(-E_a - E_0)/RT},$$

where  $E_a$  is the activation energy for adsorption of reactant, and equation (6) becomes

$$E_{\text{exp}} = (1 - \theta)(E_0 + E_a) + \theta E_p.$$

The foregoing treatment is undoubtedly over-simplified. For example  $E_p$  has been considered independent of  $T$ . Nevertheless, it is helpful towards understanding what the significance of the experimental activation energy may be, while a more rigorous treatment leads to equations which contain constants that cannot readily be found by experiment.

In the conversion of 2,3-butanediol to methyl vinyl carbinol over thoria,  $E_{\text{exp}} = 25,000$  calories per mole, and  $E_p = 14,000$  when  $\theta$  is approximately 0.5. Since  $E_1$  is 2,000 calories greater for diol than for water,  $E$  is taken to be 16,000 although at 350 °C. it is possibly larger than at 200 °C.  $E_0$  is thus estimated by equation (6) to be 51,000, compared with 27,000 obtained from equation (2), or 41,000 from the equation

$$E_{\text{exp}} = E_0 - E$$

for an unretarded reaction.

(c) *Remarks on the Adsorption of Butadiene, Diol, MEK, and MVC*

Already at 110 °C. the rate of deposition of butadiene dimer on the adsorbent was sufficient to cause serious error in the measured isotherm for adsorption of the diene. Butadiene is weakly adsorbed by thoria so that during the catalytic dehydration of the diol the degree to which liberated diene is adsorbed will be quite small. Nevertheless, some reduction in catalyst surface due to dimer formation in the gas phase must be expected.

The reduction of available surface of adsorbent that took place between one run and the next, when diol isotherms were being measured, until only about 18 per cent. was available, suggests that the proportion of adsorption sites which are active during catalysis is not small.

It seems unlikely that appreciable poisoning of the catalyst by methyl heptenone can occur during the catalytic dehydration of 2,3-butanediol at temperatures of 300 to 400 °C. Adsorption of MEK formed as a by-product will lead to only very small coverage of the catalyst surface, so that the rate of bimolecular MEK dehydration, if it takes place on the surface, will be negligible. In addition, methyl heptenone on the surface can be expected to desorb readily at the high temperature, and without decomposition except perhaps in the presence of oxygen gas.



There is not sufficient evidence to determine whether MVC is dehydrated on thoria directly to diene or by way of 2-butene-1-ol. It seems reasonable, however, to assume that the same mechanism holds for the removal of the second as for the first molecule of water, so that the indirect route to the diene is an improbable one. MVC is so readily dehydrated to butadiene that it is evidently not the product which retards the first stage dehydration of the diol.

In some unreported measurements of the rate of diol dehydration at 400 °C. it was found that the efficiency of the catalyst decreased markedly after a few minutes. This, together with the results discussed above, means that large-scale production of butadiene from 2,3-butanediol catalytically will necessitate fluidization of the catalyst, so that it can be regenerated conveniently at frequent intervals.

#### V. ACKNOWLEDGMENTS

The aid of a number of members of the Division of Industrial Chemistry, C.S.I.R.O., is acknowledged, and particularly the helpful advice of Mr. D. R. Zeidler, Dr. H. H. Hatt, Mr. K. E. Murray, and Dr. A. L. G. Rees. Mr. M. Borrell assisted with some of the adsorption measurements.

#### VI. REFERENCES

- (1) WINFIELD, M. E.—*J. Coun. Sci. Industr. Res. Aust.* **18** : 412-23 (1945).
- (2) HINSHELWOOD, C. N.—“The Kinetics of Chemical Change.” (Oxford, 1940.)
- (3) GLASTONE, S., LAIDLER, K. J., and EYRING, H.—“The Theory of Rate Processes.” (McGraw-Hill: New York, 1941.)
- (4) EMMETT, P. H., and BRUNAUER, S.—*J. Amer. Chem. Soc.* **59** : 310-5 (1937).
- (5) BLATT, A. H.—“Organic Syntheses.” *Collective*, Vol. 2, p. 102. (John Wiley & Sons: New York, 1943.)
- (6) KHOKLOVKIN, M. A., and KALACHEVA, A. V.—*Sintet. Kauchuk.* **1** : 28-30 (1936).
- (7) HARKINS, W. D., and JURA, G.—*J. Chem. Phys.* **12** : 112-3 (1944); *J. Amer. Chem. Soc.* **66** : 1366-73 (1944).
- (8) BRUNAUER, S., EMMETT, P. H., and TELLER, E.—*J. Amer. Chem. Soc.* **60** : 309 (1938).
- (9) EMMETT, P. H., and BRUNAUER, S.—*J. Amer. Chem. Soc.* **59** : 1558 (1937).
- (10) WOOD, L. A., and HIGGINS, C. F.—Letter Circ. LC710 U.S. Dep. Comm., Nat. Bur. Stand. (Dec. 9, 1942).
- (11) BRUNAUER, S.—“The Adsorption of Gases and Vapors.” Vol. 1. (Oxford, 1943.)
- (12) ABBOTT, A. E., KON, G. A. R., and SATCHELL, R. D.—*J. Chem. Soc.* **1928** : 2514-24 (1928).
- (13) FOWLER, R. H.—*Proc. Camb. Phil. Soc.* **32** : 144-51 (1936).
- (14) HAHN, O., and SENFTNER, V.—*Z. Phy. Chem. A* **170** : 191-211 (1934).
- (15) BARRER, R. M.—“Diffusion in and Through Solids.” (Cambridge, 1941.)
- (16) FOEX, M.—*C.R. Acad. Sci. Paris* **215** : 534-6 (1942).
- (17) DE BOER, J. H.—“Electron Emission and Adsorption Phenomena.” (Cambridge, 1935.)
- (18) BEECK, O., SMITH, A. E., and WHEELER, A.—*Proc. Roy. Soc. A* **177** : 62-90 (1940).
- (19) MACDONALD, J. Y.—*J. Chem. Soc.* **1936** : 839-47, 832-8 (1936).
- (20) GRIFFITH, R. L.—*J. Chem. Phys.* **14** : 408-15 (1946).
- (21) WAGNER, C., and HAUFFE, K.—*Z. Elektrochem.* **45** : 409-26 (1939).
- (22) BEECK, O.—*Rev. Mod. Phys.* **17** : 61-71 (1945).
- (23) DIXON, J. K.—*J. Amer. Chem. Soc.* **53** : 1763-73 (1931).



# LIQUID-VAPOUR EQUILIBRIA

## I. THE SYSTEMS CARBON TETRACHLORIDE-CYCLOHEXANE AND WATER-ACETIC ACID

By I. BROWN\* and A. H. EWALD\*

[*Manuscript received January 19, 1950*]

### *Summary*

An apparatus for the accurate determination of liquid-vapour equilibria is described. Data obtained with it for the system carbon tetrachloride-cyclohexane at 70 °C. and the thermodynamic functions derived from these are shown to agree with the corresponding values determined by Scatchard, Wood, and Mochel(1). Equilibrium data of the system water-acetic acid have been measured at 760 mm. Hg, particular attention being given to high acetic acid concentrations. These results are compared with published values and the excess chemical potentials of the components have been calculated.

### I. INTRODUCTION

The liquid-vapour equilibria of selected non-ideal liquid mixtures are being studied as part of an investigation of the processes of azeotropic and extractive distillation. A survey of the literature shows that the liquid-vapour equilibria of only very few systems have been measured with sufficient accuracy to yield significant values of the derived excess thermodynamic functions. These have been defined by Scatchard, Raymond, and Gilmann(2): they are required to relate the properties of liquid mixtures to those of their components.

The most suitable form of apparatus for the determination of liquid-vapour equilibria is the recirculating type of still. This provides adequate samples of both the liquid and the condensed vapour but introduces the danger of the establishment of a false state of equilibrium due to enrichment of the vapour in the more volatile component by partial condensation; alternatively the vapour may be impoverished with respect to this component by entrainment of liquid droplets. In a recent critical review Fowler(3) pointed out that these effects occur with several published still designs, and Garner(4) demonstrated that widely different results can be obtained for the same mixture by using different stills. In the present paper an improved equilibrium still is described, and data obtained with it show that this still permits the measurement of true liquid-vapour equilibrium concentrations, pressures, and temperatures.

### II. APPARATUS

#### (a) *Equilibrium Still*

The most satisfactory stills are those of Scatchard, Raymond, and Gilmann(2) and Gillespie(5) (see also 6). Gillespie (personal communication, December 1947) realized, that in his still, samples taken from the boiler do not

\* Division of Industrial Chemistry, C.S.I.R.O., Melbourne.

represent the liquid which is in equilibrium with the vapour leaving the disengagement chamber; and he suggested the inclusion of a liquid trap between the disengagement chamber and the return line to the boiler(7). In a still modified in this manner the boiler and Cottrell pump may be regarded as merely providing a mixture of liquid and vapour in equilibrium which after splashing

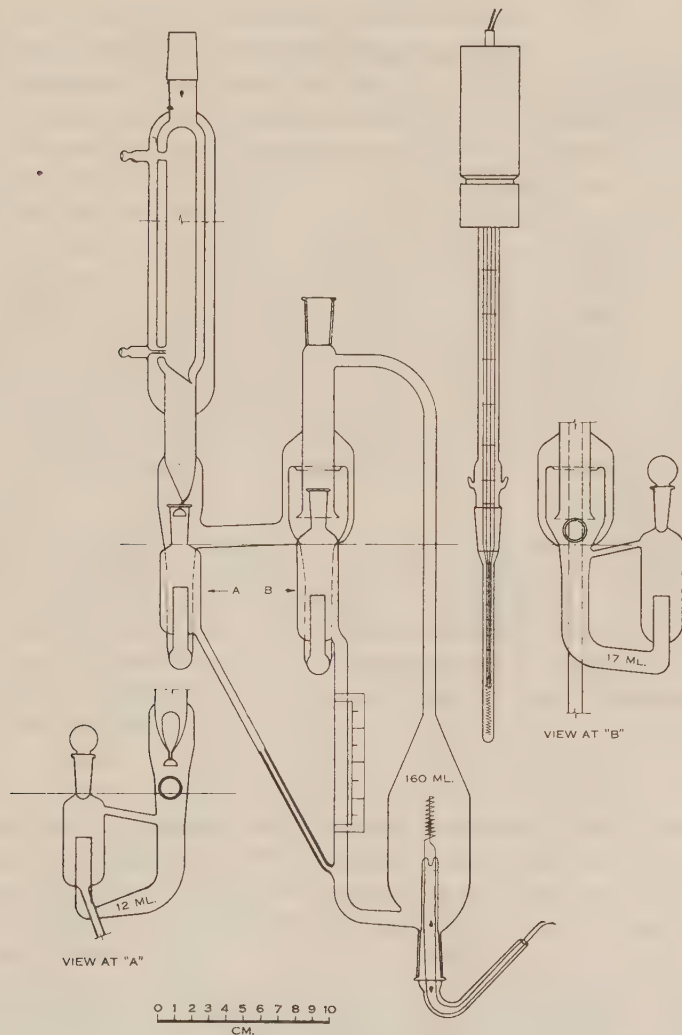


Fig. 1.—Equilibrium still.

on to the thermometer is separated in the disengagement chamber to give samples of liquid and vapour which are in equilibrium, irrespective of any condensation or pressure effects in the pump-tube.

The equilibrium still employed in this work is shown in Figure 1. The essential features of the still described by Gillespie(5) have been retained in the design, but a liquid trap has been included as suggested. A new kind of

internal heater\* was found necessary to obtain continuous nucleation and steady boiling, and the design of the boiler assembly\* had to be altered to give satisfactory pumping at pressures down to 50 mm. Hg. The disengagement chamber was enlarged to accommodate a platinum resistance thermometer and the design of the sample traps was altered to eliminate stopcocks.

### *(b) Auxiliary Equipment*

The apparatus for the control and measurement of pressure was essentially the same as that described by Scatchard, Raymond, and Gilman(2); nitrogen was used as the confining gas. The mercury levels of the precision barometer were determined by means of two telescopes and verniers, sliding on a cathetometer with a calibrated scale. The mean temperature of the barometer was determined at each reading by thermometers, mounted as close as possible to the centre of the upper and lower barometer tubes, and immersed in mercury filled tubes of the same diameter as the barometer tubes. The mercury heights which could be measured to  $\pm 0.04$  mm. were converted to standard mm. of mercury by the methods described by Beattie, Benedict, and Blaisdell(8); the pressures are correct to  $\pm 0.05$  mm. Hg. Corrections were applied for the temperature and ruling errors of the cathetometer scale, the mean temperature of the mercury, the gravitational constant at the barometer, and the capillary depressions. The corrections due to the heights of confining gas and vapour were found to be negligible.

Temperatures were determined with the specially constructed Meyer type platinum resistance thermometer shown in Figure 1, a Tinsley vernier potentiometer, and a calibrated 100 ohm standard resistance. The thermometer was calibrated at the ice and steam points and compared with a standard resistance thermometer at temperatures up to 250 °C. Its resistance at the ice point was checked periodically and remained unchanged. Temperatures were measured to  $\pm 0.003$  °C. and are recorded to  $\pm 0.01$  °C. on the international scale.

## III. OPERATION OF THE STILL

The optimum still charge was known after the initial trials and was measured by means of a level scale shown in Figure 1; a variation of 5 per cent. was permissible. Steady boiling was assured from the beginning by operating the internal heater at its maximum power input (40 watts) before connecting the still to the pressure control system. When the Cottrell pump was operating, both heaters were adjusted in accordance with the findings described subsequently in Section IV. Only small adjustments were required after the first 20 minutes of a run in order to maintain the system either at constant temperature or constant pressure.

\* The boiler was wound with 4.6 m. of 0.405 mm. nichrome wire over a layer of asbestos tape. The wire was set in alundum cement and lagged with asbestos cloth and an outer cover of aluminium foil. The internal heater consisted of a coil of 25 cm. of 0.32 mm. platinum wire, with a short length of thick-walled soda-glass capillary over the central wire sealed to it at the lower end. Both heaters were controlled by variable voltage transformers. The Cottrell pump-tube and disengagement chamber were lagged with asbestos tape over a layer of aluminium foil.

On turning off the external heater in preparation for sampling, its large heat capacity caused boiling to continue for 1 to 2 minutes and, to avoid bumping at this stage, the internal heater was switched off only when the pumping in the Cottrell tube had almost ceased. The still was then vented to atmosphere through a drying tube and, when working at 760 mm. Hg, any subsequent distillation of vapour through the Cottrell tube was prevented by spraying water on the lagging of the tube. Samples were then withdrawn from the two traps.

When adjusting the composition of the still charge to that required for the next run, the volume of mixture to be removed, and the amount of pure component to be added were calculated, and only occasionally was a subsequent adjustment of the liquid level necessary.

#### IV. STILL PERFORMANCE TESTS

(i) *Boiling*.—Test runs with water and aqueous solutions of ethanol and acetic acid showed that the internal heater gave adequate nucleation and smooth boiling at pressures from atmospheric to as low as 50 mm. Hg.

(ii) *Distillation and Liquid Flow-Rates*.—The distillation rate depended mainly on the external heating, while the rate at which the liquid circulated through the Cottrell tube was controlled by the internal heater and liquid level. Flow-rates through the liquid trap were measured, while the still was operating, by using a calibrated pump to force liquid from the trap to the return line at a rate sufficient to maintain a constant level in the trap. Distillation rates were obtained from the drop counter in the still. For dilute aqueous solutions liquid flow-rates were up to 430 ml./min. and distillation rates up to 3.8 ml./min., while with acetone the corresponding figures were 200 and 16.5 ml./min.

(iii) *Entrainment*.—The absence of entrainment at high liquid and vapour flow-rates was confirmed by the following tests.

The boiler and the liquid trap were filled with a 1 per cent. solution of fluorescein in ethanol, and on operating the still at distillation rates up to 18 ml./min. and liquid flow-rates up to 400 ml./min., no colour could be detected in the vapour trap after 9 hours. A similar test with aqueous potassium permanganate also indicated no entrainment.

Furthermore, vapour samples from two runs with a 75 mol. per cent. solution of acetic acid in water, at distillation rates of 3.3 and 13.2 ml./min. respectively, had identical concentrations.

(iv) *Distillation in the Vapour Trap*.—Distillation from the liquid in the trap on to the walls of the vapour trap could conceivably cause impoverishment of the condensate sample. During one run with aqueous acetic acid the walls of the vapour trap were heated electrically to a temperature just above that of the condensing vapour, but this had no effect on the composition of the vapour sample.

(v) *Time Required to Reach Equilibrium*.—Measurement of the boiling temperature to  $\pm 0.005^\circ\text{C}$ . is not sufficiently sensitive to indicate the final approach to equilibrium; two mixtures of widely different relative volatility ( $\alpha$ )



were used in the following manner to determine the time required to reach equilibrium in the still.

Aqueous acetic acid (30 wt. %,  $\alpha=1.43$ ) was used in six runs lasting from 0.25 to 2.5 hours at a distillation rate of 7 ml./min., and samples from these were analysed with an accuracy of  $\pm 0.23$  per cent. by the methods described later. After 2 hours no further change in the composition was found.

The second test mixture was a 30 mol. per cent. solution of nitrobenzene in cyclohexane ( $\alpha=30.4$ ) which was distilled at a rate of 15 ml./min. Samples of 0.25 ml. were taken from the vapour trap at 3 minute intervals and analysed refractometrically with an accuracy of  $\pm 0.05$  per cent. No change in composition was found after 13 minutes.

These findings are in marked contrast with the statements of Othmer(9) and Fowler(10) that 41.5 hours are required to reach equilibrium in a Gillespie still, but are in agreement with the findings of Rieder and Thompson(11) who found that 3 hours were adequate to give equilibrium for the system ethanol-water in such a still. The time required to reach equilibrium in any recirculating still will depend both on the ratio of the distillation rate to the volume of liquid in the vapour trap, and on the relative volatility of the mixture; and will be small for high values of these ratios. The liquid flow-rate, which is always far greater than the distillation rate, will in no case be a controlling factor.

(vi) *Conclusion.*—The following conclusions were reached with regard to the operation of the still:

(1) The liquid level should be as low as is consistent with adequate pumping. Satisfactory operation was obtained with a total charge ranging from 182 to 193 ml.

(2) The internal heating should be as low as is consistent with even boiling and adequate wetting of the thermometer (2 to 20 watts).

(3) To obtain a high distillation rate the external heating should be as high as possible without causing surging of the liquid (80 to 300 watts).

## V. THE SYSTEM CARBON TETRA CHLORIDE-CYCLOHEXANE

### (a) *Purification of Components*

Commercial carbon tetrachloride was shaken with alcoholic potassium hydroxide, with 98 per cent. sulphuric acid, with an alkaline solution of potassium permanganate; and the product finally dried over anhydrous calcium chloride, and distilled at 760 mm. Hg through a column of 25 mm. diameter, 1.5 m. long, and filled with 3 mm. single turn glass helices. A head fraction of 60 ml. was removed using a total reflux trap as described by Scatchard, Wood, and Mochel(12), and then a fraction boiling between 76.63 and 76.65 °C. (corr.) was distilled at a reflux ratio of 50 : 1. The boiling point, as determined in the equilibrium still, and the density of the product are compared with published values in Table 1.

cycloHexane (B.D.H.) was nitrated following the method of Scatchard, Wood, and Mochel(12) to remove aromatic impurities; was dried over anhydrous magnesium sulphate; and was fractionated at 760 mm. Hg at a reflux ratio of

100:1 in the column previously described. The fraction boiling at 80.70 to 80.71 °C. (corr.) was collected as product and no aromatic impurities were revealed by its ultraviolet absorption spectrum. The boiling point of the product, as determined in the equilibrium still, its freezing point, as found from a cooling curve, and its density are compared with published values in Table 1.

TABLE 1  
COMPONENT PROPERTIES

	Carbon Tetrachloride		<i>cyclo</i> Hexane		
	Boiling Point (°C.)	$d_4^{25}$	Boiling Point (°C.)	Freezing Point (°C.)	$d_4^{25}$
This investigation .. ..	76.66	1.58429	80.72	$6.54 \pm 0.02$	0.77390
Scatchard, Wood, and Mochel(1)	76.687	1.58414	80.739	6.49	0.77383
Grimm(13) .. ..	76.71	1.58413	—	—	—
Rossini <i>et al.</i> (14) .. ..	—	—	80.738	6.55	0.77389

### (b) Experimental Procedure

(i) *Analysis*.—Liquid compositions were determined from density measurements at  $25.00 \pm 0.01$  °C. Seven mixtures of the components were prepared by weighing in narrow-necked, stoppered 25 ml. flasks, allowance being made for the loss of first component in the air-vapour mixture displaced by the addition of the second component through a stainless steel cannula(15). The densities of the mixtures were determined in duplicate in single stem 10 ml. pycnometers fitted with glass caps. The meniscus level in the calibrated capillary stem (1.4 mm. diameter) was measured to  $\pm 0.02$  mm. with a telemicroscope mounted on a cathetometer.

The measured densities were related to the compositions by equations (1) and (2) as suggested by Scatchard(16).

$$1/\bar{d} = 1/\bar{d}_0 + \Delta, \quad \dots\dots\dots (1)$$

$$\Delta = A.x_1x_2 + B.x_1x_2(x_1 - x_2), \quad \dots\dots\dots (2)$$

where  $1/\bar{d}_0$  is the specific volume, assuming the volume per mole to be additive,  $\Delta$  is the deviation from this value, and  $x_1$  and  $x_2$  are the mol. fractions of carbon tetrachloride and *cyclo*hexane respectively. Values of  $\Delta$  were determined from the densities and calculated compositions of the synthetic mixtures shown in Table 2 and the best values of the constants in equation (2) were found by the method of least squares to be:  $A = +0.00479$ ,  $B = -0.00107$ .

The last column in Table 2 shows the differences between the densities calculated from equations (1) and (2) and the experimental values, and shows these to agree within the experimental probable error of  $\pm 0.00004$  found from quadruplicate density measurements of each of the pure components.

TABLE 2  
COMPOSITIONS AND DENSITIES AT 25 °C. OF SYNTHETIC MIXTURES OF CCl<sub>4</sub> AND C<sub>6</sub>H<sub>12</sub>

Mol. Fraction CCl <sub>4</sub>	$d_4^{25}$ exp.	$\Delta$	$d_{\text{calc.}} - d_{\text{exp.}}$
0	0.77390	0	—
0.08969	0.83904	+0.00046	—0.00003
0.25593	0.96339	0.00101	—0.00001
0.39523	1.07117	0.00120	—
0.51572	1.16730	0.00119	+0.00008
0.64799	1.27611	0.00102	—
0.77261	1.38191	0.00074	+0.00007
0.88927	1.48426	0.00039	—0.00001
1	1.58429	0	—

The values of  $\Delta$  calculated by equation (2) and shown in Table 2 were plotted on a large scale *versus* density, and this graph was used to find  $d_0$  for each mixture; the composition was then calculated from the expression.

$$x_1 = \frac{d_0 V_2 - M_2}{(M_1 - M_2) - (V_1 - V_2)d_0}, \quad \dots\dots\dots (3)$$

where  $x$ ,  $M$ ,  $V$  are mol. fraction, molecular weight, and molecular volume respectively, and where subscript 1 refers to carbon tetrachloride and subscript 2 to cyclohexane. The experimental error in these compositions was calculated to be  $\pm 0.00014$  by assuming the error in  $d_0$  to be  $\pm 0.00005$ .

(ii) *Operation.*—The liquid and vapour equilibrium compositions of 12 mixtures were determined at 70 °C., and the vapour pressures of the pure components were measured in the still at 10 °C. intervals between 30 and 80 °C. The still was operated as previously described, with the pressure controlled to maintain a boiling temperature of  $70.00 \pm 0.01$  °C. A liquid throughput of 200–300 ml./min. and a distillation rate of about 16 ml./min. were used, and the still was run for 3 to 4 hours before sampling. The time required to reach equilibrium was estimated to be 1.5 to 2 hours for this system, by taking into consideration the value of  $\alpha$  (1.13) and the distillation rate, as compared with those used in the tests described in Section IV (v).

Samples were taken by means of glass syringes fitted with stainless steel cannulae to permit direct injection into the pycnometers. The cannulae were inserted into the traps through rubber serum caps and were provided with small water-jackets through which ice-water was circulated. A sampling and injection rate of 12 ml./min. was sufficiently slow to cool the samples to below 25 °C. After the pycnometers had been placed in the constant temperature bath, the liquid level in the stem was brought close to the zero mark in one operation by withdrawing liquid through a fine capillary.

(c) *Results*

The vapour pressures of the pure components are shown in Table 3 and are consistently about 0.25 mm. higher than those given by Scatchard. This corresponds to the experimental error of  $\pm 0.01^\circ\text{C}$ . in the measurement of temperature.

TABLE 3  
VAPOUR PRESSURES OF PURE COMPONENTS

Temperature ( $^\circ\text{C}$ .)	Pressure (mm. Hg) (Standard)	
	Carbon Tetrachloride	<i>cyclo</i> Hexane
80.72	—	760.00
76.66	760.00	—
70.00	617.43	544.23
60.00	444.44	389.60
50.00	312.23	272.02
40.00	213.42	184.75
30.00	—	121.91

TABLE 4  
CARBON TETRACHLORIDE-CYCLOHEXANE AT  $70^\circ\text{C}$ .

Mol. Fraction $\text{CCl}_4$		Pressure (mm. Hg)	$\mu_1^E$ (cal./mol.)	$\mu_2^E$ (cal./mol.)	$G_x^E$ (cal./mol.)	$G_x^E$ exp. $-G_x^E$ calc.
Liquid $x_1$	Vapour $y_1$					
0.0538	0.0641	551.10	45.20	0.76	3.15	+0.47
0.1396	0.1622	560.58	38.54	1.37	6.56	+0.15
0.2327	0.2629	570.28	30.78	3.44	9.80	+0.13
0.2641	0.2968	573.20	30.43	3.24	10.42	-0.15
0.3135	0.3476	577.94	26.88	4.83	11.74	-0.06
0.4041	0.4388	585.90	21.49	7.77	13.32	-0.06
0.5153	0.5464	594.31	14.80	12.74	13.81	-0.29
0.5994	0.6259	600.32	10.92	17.98	13.75	+0.03
0.7012	0.7209	606.27	6.84	24.60	12.15	+0.00
0.8009	0.8125	611.19	3.06	35.59	9.54	+0.17
0.8995	0.9045	614.95	1.12	45.51	5.58	+0.19
0.9479	0.9500	616.17	0.17	53.18	2.93	-0.03

The liquid and vapour mol. fractions  $x_1$  and  $y_1$  and the total vapour pressure  $P$  of the mixture are shown in Table 4. The excess chemical potentials  $\mu_1^E$



and  $\mu_2^E$  and the excess free energy  $G_x^E$ , which are the excess of these functions over their values in an ideal solution, are shown in columns 4, 5, and 6 and were calculated by the following equations:

$$\mu_1^E = RT \ln \gamma_1 = RT \ln \frac{Py_1}{P_1x_1} + C(V_1 - \beta_1)(P_1 - P), \dots \dots \dots (4)$$

$$\mu_2^E = RT \ln \gamma_2 = RT \ln \frac{Py_2}{P_2x_2} + C(V_2 - \beta_2)(P_2 - P), \dots \dots \dots (5)$$

$$G_x^E = x_1\mu_1^E + x_2\mu_2^E \dots \dots \dots (6)$$

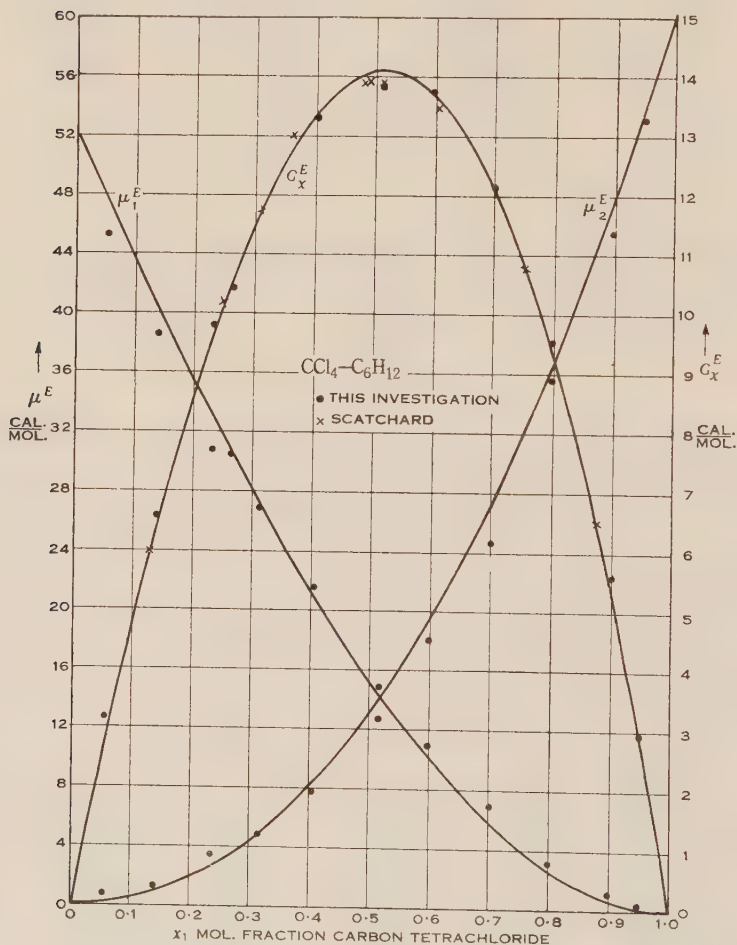


Fig. 2

The last term in equations (4) and (5) is a correction for the non-ideality of the vapour phase.

In these equations  $P$  is the total pressure in mm. Hg;  $P_1, P_2, V_1, V_2$  are the vapour pressures and liquid molar volumes of the pure components at 70 °C.;  $\gamma_1$  and  $\gamma_2$  are the activity coefficients in the mixtures; and  $\beta_1$  and  $\beta_2$  are the

second virial coefficients in the equations of state for the pure components. The constant  $C$  converts mm. Hg-litres to calories ( $C=0.0318657$ ). The molar volumes of the pure liquids at 70 °C. were determined from published densities.

$$V_1=0.10276 \text{ litre(17)}$$

$$V_2=0.11519 \text{ litre(18, 19, 20).}$$

The virial coefficients were calculated from critical data by the Berthelot equation and the values,  $\beta_1=-1.054$  litres and  $\beta_2=-1.154$  litres, are in agreement with the experimental values found by Lambert *et al.*(21).

A three suffix Margules equation in the form given by Wohl(22) was fitted to the calculated values of  $G_x^E$  and the following equations obtained for the excess thermodynamic functions:

$$G_x^E = x_1 x_2 (60.3978 x_1 + 52.2509 x_2), \dots \dots \dots (7)$$

$$\mu_1^E = x_2^2 (52.2509 + 16.2938 x_1), \dots \dots \dots (8)$$

$$\mu_2^E = x_1^2 (60.3978 - 16.2938 x_2) \dots \dots \dots (9)$$

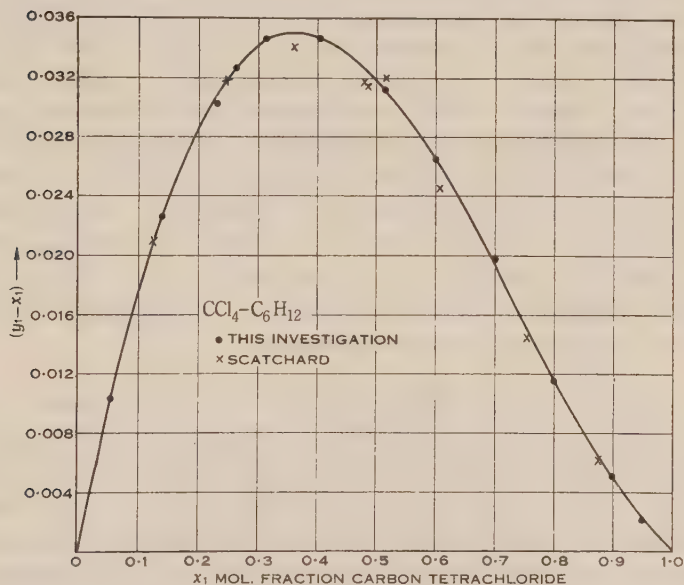


Fig. 3

Column 7 in Table 4 shows the differences between  $G_x^E$  calculated from equation (7) and from the experimental values. These are within the experimental error in  $G_x^E$  which was estimated to be  $\pm 1.2$  cal./mol.

A comparison of these data with those obtained by Scatchard, Wood, and Mochel(1) is shown in Figure 2 in which the curves represent values of  $G_x^E$ ,  $\mu_1^E$  and  $\mu_2^E$  calculated from equations (7), (8), and (9), and the points shown are values calculated from experimental data obtained in this investigation and by Scatchard. In Figure 3 these experimental data are plotted as  $(y_1 - x_1)$  against  $x_1$ .

The internal consistency of the data was established by applying the Gibbs Duhem equation in two ways. Equation (7) is an integrated form of this equation and was shown to represent the experimental values of  $G_x^E$  within the estimated experimental error; the values must thus be consistent among themselves and must represent the thermodynamically defined quantities. The second check was obtained by the method of Redlich and Kister(23). The experimental values of  $RT \ln \gamma_1$  and  $RT \ln \gamma_2$  were plotted separately against  $x_1$  and the difference of the areas under the two curves was found. This amounted to 4.5 per cent., but it was estimated that experimental errors in  $RT \ln \gamma_1$  and  $RT \ln \gamma_2$  could cause a difference of about 18 per cent. in these areas.

## VI. THE SYSTEM WATER-ACETIC ACID AT 760 MM. Hg

### (a) Purification of Components

One component was freshly prepared conductivity water.

The other component, acetic acid obtained by hydrolysis of purified acetic anhydride (B.D.H., A.R.), was refluxed with 2 per cent. by weight of solid potassium permanganate as recommended by Bousfield and Lowry(24) and distilled from the residue. It was then fractionated at a reflux ratio of 40 : 1 through a 50 cm. column similar to that used for the carbon tetrachloride purification. The water content of the product was found by the Karl Fischer method to be 0.006 per cent. by weight, and the acid was finally dried by refluxing with 100 per cent. excess of purified acetic anhydride,\* as recommended by Orton and Bradfield(25), but using concentrated sulphuric acid as catalyst.

The acid was then carefully fractionated through the 50 cm. column and a middle portion collected in a flask from which it was directly distilled into the equilibrium still and freezing point apparatus.

### (b) The Normal Boiling Point of Pure Acetic Acid

The wide variation in the reported boiling points of even carefully purified samples of acetic acid (117.8–118.5 °C.) is due to the extreme difficulty of drying it. The only two determinations found, in which the same sample was used to determine both the boiling point and the freezing point, were those of Bousfield and Lowry(24) and Kendall and Brakeley(26). Their results, together with four values obtained in this investigation, are shown in Table 5.

Figure 4 shows a plot of boiling point *versus* freezing point of the samples of Table 5 together with their claimed accuracies. The most reliable values of the freezing point were considered to be those of de Visser(28) (16.5965 °C.) and MacDougall(29) (16.612, 16.616 °C.) and, assuming the latter to be correct, a straight line drawn through all the values shows the normal boiling point to

\* The acetic anhydride used in the final dehydration was obtained from fractionated B.D.H., A.R. anhydride which had been refluxed with carefully purified chromic anhydride, filtered, and refractionated. Since  $\text{CrO}_3$  containing traces of water and sulphuric acid reacts with acetic anhydride with explosive violence, the chromic anhydride (B.D.H., A.R.) was recrystallized from water and dried at 110 °C.

TABLE 5  
THE FREEZING POINT AND BOILING POINT OF ACETIC ACID

Sample	Boiling Point/760 (°C.)	Freezing Point (°C.)	Wt. % H <sub>2</sub> O
1 .. .. .	117.85 ± 0.01	16.56 ± 0.02	0.0242*
2 .. .. .	117.90	16.60*	0.00742
3 .. .. .	117.92	16.60*	0.00064
4 .. .. .	117.95	16.60*	0.00753
Bousfield and Lowry ..	117.88 ± 0.05	16.60 ± 0.005	—
Kendall and Brakeley ..	117.8 ± 0.1	16.57 ± 0.05	—

\* These values were determined from the cryoscopic constant for acetic acid 39.0°/g. mole(27) and the value of MacDougall (16.614 °C.) for the freezing point of the pure acid.

be 117.95 ± 0.01 °C. This agrees within ± 0.01 °C. with the value obtained in this investigation by extrapolation of a plot of liquid and vapour mol. fractions against boiling temperature.

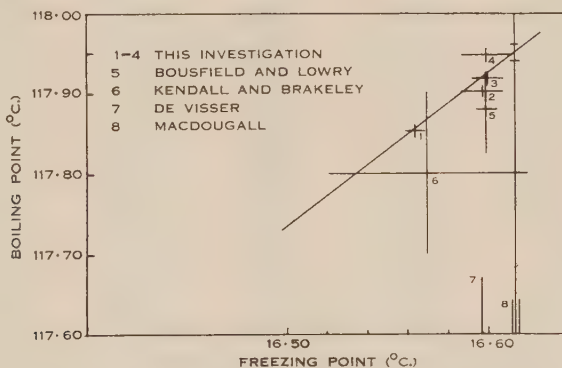


Fig. 4.—Acetic acid: boiling point and freezing point.

### (c) Experimental Procedure

The still was operated for 3 to 4 hours at distillation rates of up to 13.2 ml./min. and liquid flow-rates of 200 to 300 ml./min. while the pressure was controlled at one standard atmosphere to within ± 0.2 mm. Hg.

The hygroscopic nature of concentrated acetic acid made it necessary to exclude contact with the atmosphere during sampling, and two syphons as shown in Figure 5 were used to withdraw samples of the required size by applying suction at *A* with a graduated glass syringe. *B* is a rubber stopper which closed the trap during sampling.

Mixtures containing less than 0.33 mol. fraction of water were sampled into dry, 100 ml. conical flasks which had been weighed and cooled. The flasks were stoppered and re-weighed after attaining room temperature, and the water in the samples estimated by the Karl Fischer method using an electro-metric end point. The water equivalents of the Karl Fischer reagent and the



aqueous methanol used for back titration were determined at the beginning, the middle, and the end of each series of titrations by means of water weighed out of a small weight pipette. Two empty flasks were titrated with each set of samples and allowance was made for these blank titres.

Samples containing a higher concentration of water were withdrawn into calibrated 100 ml. volumetric flasks, which had been half filled with cold distilled water and weighed. The weight of the samples was found by re-weighing the flasks; and triplicate aliquots were titrated with decinormal barium hydroxide to a phenolphthalein end point. One sample was analysed by both this and the Karl Fischer method and the results were in agreement.

The analytical errors were estimated as follows:

(i) *Karl Fischer Method*.—Mixtures having mol. fractions of water between 0.33 and 0.1 were sampled and analysed in duplicate and the mean error was found to be  $\pm 0.16$  per cent. which corresponds to  $\pm 0.0003$  in mol. fraction. Only single samples of adequate size could be taken of mixtures containing less

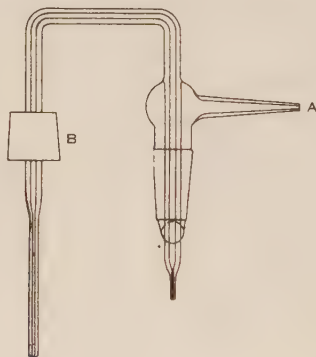


Fig. 5.—Sampling syphon.

than 0.1 mol. fraction of water, and the analytical errors were estimated to be  $\pm 1$  per cent. ( $\pm 0.0001$  in mol. fraction) in the range from 0.1 to 0.003 mol. fraction. At mol. fractions below 0.0004, even the largest sample that could be taken gave a titre which was comparable in size with the titre of the blank flasks, and the errors in this range were estimated to be  $\pm 40$  per cent. which corresponds to  $\pm 0.0001$  in mol. fraction.

(ii) *Barium Hydroxide Method*.—The errors were determined from analyses of duplicate samples, each titrated in triplicate, and were found to range from  $\pm 0.15$  per cent. at a mol. fraction of water of 0.5 to  $\pm 0.06$  per cent. at a mol. fraction of 0.99.

#### (d) Results

The results of 19 determinations of the liquid-vapour equilibrium compositions and temperatures for this system are shown in Table 6, together with the excess chemical potentials ( $\mu^E$ ) calculated from the experimental data by equations (4) and (5). The additional data used in these calculations were:

*Water.*—The molar volumes were calculated from published density values(30); the vapour pressures  $P_1$  and virial coefficients  $\beta_1$  were obtained from the equation of state of Keyes, Smith, and Gerry(31).

TABLE 6  
WATER-ACETIC ACID AT 760 MM. Hg

Mol. Fraction of Water		Temperature (°C.)	Cal./Mol. (uncorr.)		Cal./Mol. (corr.)	
Liquid $x_1$	Vapour $y_1$		$\mu_1^E$	$\mu_2^E$	$\mu_1^E$	$\mu_2^E$
0.000210	0.000226	117.96	—414.8	—7.9	—406.6	—7.9
0.000213	0.000347	117.92	—91.9	+0.1	—83.8	+0.0
0.000247	0.000450	117.91	+4.8	+0.5	+3.4	+0.3
0.00340	0.00689	117.64	+85.0	+4.2	+93.0	+1.5
0.00545	0.01122	117.51	100.0	+4.2	108.1	+0.8
0.04741	0.09790	115.03	164.6	+27.4	170.5	+1.1
0.08120	0.1446	113.81	80.1	43.3	86.4	+6.1
0.1497	0.2382	111.51	51.8	69.8	57.0	+12.2
0.2198	0.3273	109.84	44.0	80.7	48.4	+14.4
0.2917	0.4071	108.16	37.3	99.0	40.9	+12.5
0.3378	0.4573	107.36	34.3	101.9	37.6	+8.1
0.4198	0.5496	105.85	48.5	98.4	51.1	—8.1
0.5359	0.6591	104.17	44.9	98.2	46.7	—22.3
0.6463	0.7524	102.86	37.8	94.1	39.1	—37.4
0.7388	0.8217	101.92	28.2	98.5	29.0	—40.8
0.8251	0.8783	101.24	14.1	129.7	14.6	—21.4
0.9210	0.9429	100.54	3.0	175.4	3.2	+21.0
0.9676	0.9761	100.24	0.1	200.0	0.2	+45.7
0.9891	0.9921	100.07	0.3	188.4	0.4	+33.6

*Acetic Acid.*—The molar volumes were calculated from density values of Young(32). Vapour pressures, measured in the equilibrium still over the range 100 to 118 °C., were used to determine the constants in the following equation :

$$\log P_2 \text{ (mm. Hg)} = 8.11401 - 2046.74 \cdot \frac{1}{T} \text{ (°K.)}.$$

Values calculated from this equation agree with the values given by Stull(33).

The virial coefficients of acetic acid  $\beta_2$  were calculated from those of water by the principle of corresponding states. The value of  $\beta_2$  at the normal boiling point of acetic acid was determined from the vapour density of 0.00316 g./ml.(34). The ratio of  $\beta_2$  to the virial coefficient of water at the same reduced temperature was found to be 42.798 and was assumed to remain constant from 100 to 118 °C. Values of  $\beta_2$  calculated on this assumption ranged from —14.54 l./mol. at 100 °C. to —12.24 l./mol. at 118 °C., and are of the same magnitude as values obtained by extrapolating the experimental vapour density data of Ritter and Simons(35) (—13.87 to —12.88 l./mol.).

In Figure 6, values of the excess chemical potentials with and without the corrections for the non-ideality of the vapours are plotted against concentration. The calculated values of  $\mu_2^E$  are somewhat indefinite due to the magnitude and the uncertainty of the virial coefficient of acetic acid. The Gibbs Duhem equation could not be used to test the consistency of the data due to this uncertainty and because the measurements were not made at constant temperature. The error in  $\mu_1^E$  at  $x_1=0.003$  is  $\pm 15$  cal./mol. or 15 per cent., but falls to  $\pm 2.5$  cal./mol. (1.5 per cent.) at  $x_1=0.05$  and the maximum in the

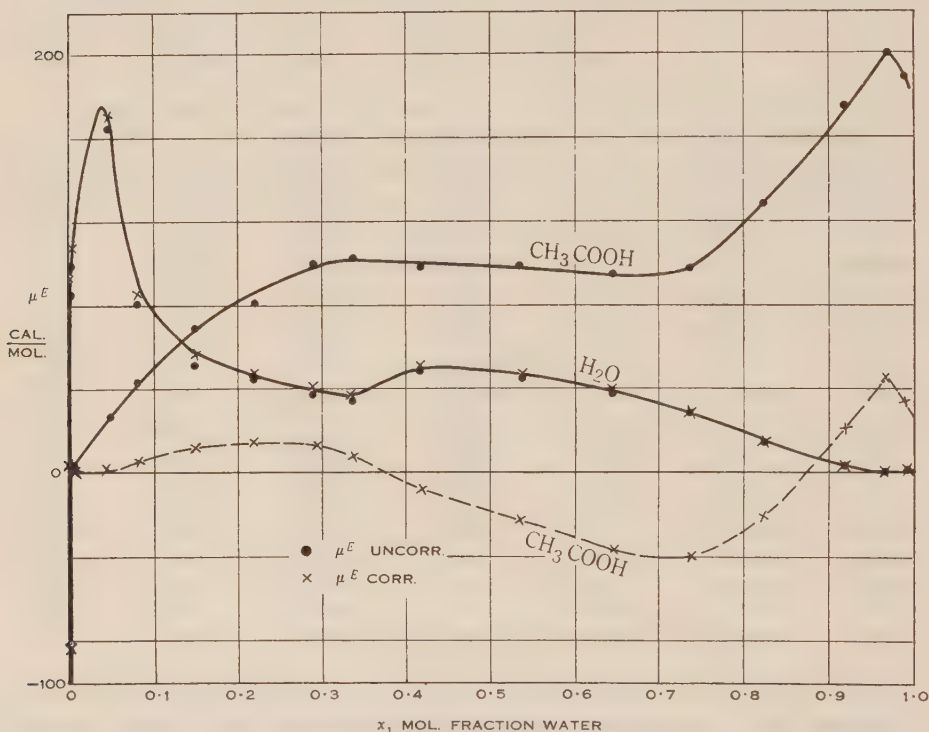


Fig. 6

$\mu_1^E$  curve at this concentration is therefore based on experimentally significant values. At water concentrations below 0.003 mol. fraction the experimental error is large and the values of  $\mu_1^E$  uncertain; however, the negative values obtained at the two lowest water concentrations may be significant as they lie on a smooth curve drawn through them and the three succeeding positive values. These negative values of  $\mu_1^E$  also become plausible if one considers the very strong interaction between water and acetic acid molecules (cf. the very hygroscopic nature of concentrated acetic acid) while the maximum in the curve could indicate that superimposed on this interaction, which leads to a decrease in the free energy of the water molecules, there is the effect of the dissociation of the molecular complexes of the water, giving an increased free energy.

## (e) Comparison of Results with Other Measurements

Figure 7 shows a plot of  $(y_1 - x_1)$  versus  $x_1$  in which the curve represents the results obtained in this investigation. The results of Cornell and Montonna(36) and of York and Holmes(37) lie close to the curve while those of Povarnin and Markov(38), Othmer and Gilmont(39), and Pascal, Dupuy, and Garnier(40) show considerably higher values. The results of other workers, which have been included in a similar comparison by Othmer and Gilmont(39) have been omitted to simplify the figure. The high values of  $(y_1 - x_1)$  obtained by some

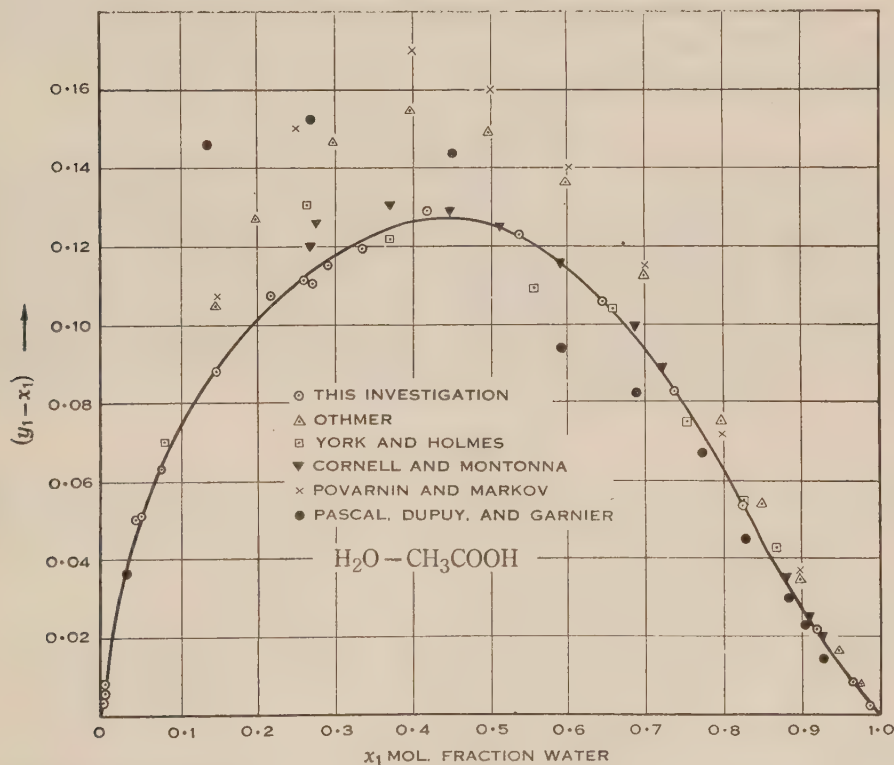


Fig. 7

workers may be due to partial condensation giving enrichment of the vapour, while low values may be due to entrainment. It has already been shown that no entrainment occurs in the apparatus described here, even at distillation rates considerably higher than those used in the actual determinations; while the design of the disengagement chamber eliminates the possibility of enrichment.

## VII. CONCLUSIONS

The liquid-vapour equilibrium still described in this paper has been shown to be free of the faults found in a number of other stills; it was found to operate satisfactorily over a wide range of conditions, and to be suitable for the determination of precise equilibrium data both at constant pressure and at constant



temperature. It has been shown that equilibrium is established in the still after approximately 2 hours of operation and that the contents of the liquid and vapour traps then have concentrations corresponding to a true liquid-vapour equilibrium.

The normal boiling point of acetic acid has been redetermined and an interpretation has been given for the variation with composition of the excess chemical potential of water in acetic acid solution.

The equilibrium still is now being used to investigate a series of binary mixtures of *n*-heptane with other non-polar components. The results obtained will be used to determine the relation between the thermodynamic properties of these mixtures and the physical properties of their components, and will be reported in a subsequent paper.

#### VIII. ACKNOWLEDGMENTS

The authors thank the members of the instrument workshops, Division of Industrial Chemistry, C.S.I.R.O., for their assistance in the construction of apparatus and, in particular, Mr. K. A. Ophel for the construction of the resistance thermometer. They wish to thank Mr. D. T. C. Gillespie for valuable suggestions, and also members of the Chemical Physics Section of the abovementioned Division for spectroscopic measurements, and the staff of the Defence Research Laboratories, and the National Standards Laboratory, C.S.I.R.O., for their advice and for the calibration of instruments.

#### IX. REFERENCES

- (1) SCATCHARD, G., WOOD, S. E., and MOCHEL, J. M.—*J. Amer. Chem. Soc.* **61** : 3206 (1939).
- (2) SCATCHARD, G., RAYMOND, C. L., and GILMANN, H. H.—*J. Amer. Chem. Soc.* **60** : 1275, 1278 (1938).
- (3) FOWLER, R. T.—*Industr. Chem. Chem. Mfr.* **24** : 717, 824 (1948).
- (4) GARNER, F. H.—*Industr. Chem. Chem. Mfr.* **25** : 238 (1949).
- (5) GILLESPIE, D. T. C.—*Industr. Engng. Chem.* (Anal. Ed.) **18** : 575 (1946).
- (6) COULSON, E. A., HALES, J. L., and HERINGTON, E. F. G.—*Trans. Faraday Soc.* **44** : 636 (1948).
- (7) FOWLER, R. T.—*J. Soc. Chem. Ind.* **68** : 131 (1949).
- (8) BEATTIE, J. A., BENEDICT, M., and BLAISDELL, B. E.—*Proc. Amer. Acad. Arts Sci.* **71** : 327 (1938).
- (9) OTHMER, D. F.—*Anal. Chem.* **20** : 763 (1948).
- (10) FOWLER, R. T.—*Chemistry and Industry*, p. 219 (1949).
- (11) RIEDER, R. M., and THOMPSON, A. R.—*Industr. Engng. Chem.* **41** : 2905 (1949).
- (12) SCATCHARD, G., WOOD, S. E., and MOCHEL, J. M.—*J. Phys. Chem.* **43** : 119 (1939).
- (13) GRIMM, H. G.—*Z. Phys. Chem.* B **2** : 181 (1929).
- (14) ROSSINI, D., *et al.*—Selected values of properties of hydrocarbons. *Circ. U.S. Nat. Bur. Stand. (Dep. Comm.)* C461 (1947).
- (15) SCATCHARD, G., WOOD, S. E., and MOCHEL, J. M.—*J. Amer. Chem. Soc.* **68** : 1957 (1946).
- (16) SCATCHARD, G.—*Chem. Rev.* **44** : 7 (1949).
- (17) YOUNG, S.—*J. Chem. Soc.* **59** : 37 (1891).
- (18) MASSART, L.—*Bull. Soc. Chim. Belg.* **45** : 83 (1936).
- (19) ROTINJANZ, L., and NAGORNOW, N.—*Z. Phys. Chem.* A **169** : 20 (1934).

- (20) YOUNG, S., and FORTEY, E. C.—*J. Chem. Soc.* **75** : 873 (1899).
- (21) LAMBERT, J. D., ROBERTS, G. A. H., ROWLINSON, J. S., and WILKINSON, V. J.—*Proc. Roy. Soc. A* **196** : 113 (1949).
- (22) WOHL, K.—*Trans. Amer. Inst. Chem. Engrs.* **42** : 215 (1946).
- (23) REDLICH, O., and KISTER, A. T.—*Industr. Engng. Chem.* **40** : 345 (1948).
- (24) BOUSFIELD, W. R., and LOWRY, T. M.—*J. Chem. Soc.* **99** : 1432 (1911).
- (25) ORTON, K. J. P., and BRADFIELD, A. E.—*J. Chem. Soc.* **1927** : 983 (1927).
- (26) KENDALL, J., and BRAKELEY, E.—*J. Amer. Chem. Soc.* **43** : 1826 (1921).
- (27) INTERNATIONAL CRITICAL TABLES **4** : 183 (1928).
- (28) DE VISSER.—*Rec. Trav. Chim.* **12** : 101, 154 (1893).
- (29) MACDOUGALL, F. H.—*J. Amer. Chem. Soc.* **58** : 2585 (1936).
- (30) INTERNATIONAL CRITICAL TABLES **3** : 25 (1928).
- (31) KEYES, F. G., SMITH, L. B., and GERRY, H. T.—*Proc. Amer. Acad. Arts Sci.* **70** : 319 (1936).
- (32) YOUNG, S.—*Sci. Proc. R. Dublin Soc.* **12** : 374 (1910).
- (33) STULL, D. R.—*Industr. Engng. Chem.* **39** : 517 (1947).
- (34) WRIGHT, R.—*J. Phys. Chem.* **36** : 2793 (1932).
- (35) RITTER, H. L., and SIMONS, J. H.—*J. Amer. Chem. Soc.* **67** : 757 (1945).
- (36) CORNELL, L. W., and MONTONNA, R. E.—*Industr. Engng. Chem.* **25** : 1331 (1933).
- (37) YORK, R., and HOLMES, R. C.—*Industr. Engng. Chem.* **34** : 345 (1942).
- (38) POVARNIN, G., and MARKOV, V.—International Critical Tables **3** : 310 (1928).
- (39) OTHMER, D. F., and GILMONT, R.—*Industr. Engng. Chem.* **36** : 1061 (1944).
- (40) PASCAL, P., DUPUY, E., and GARNIER.—*Bull. Soc. Chim.* **29** : 9 (1921).

# THE SYNTHESIS OF POLYMERS IN REDUCED WOOL

By M. LIPSON\* and R. J. HOPE\*

[Manuscript received February 17, 1950]

## Summary

From experiments *in vitro* it is found that cysteine, in the presence of certain oxidizing agents, can act as a powerful catalyst for polymerizing ethylenic monomers. This aids in explaining the reactions which occur when wool is reduced with sodium bisulphite and then treated with a monomer. Polymerization within the fibre is probably initiated by free radicals produced from reduced cystine by the action of sulphoxides or sulphones present in the wool as a result of atmospheric oxidation. Fraction (A+B) of the cystine appears to be involved as pretreatment with alkali, which converts this fraction to lanthionine, prevents polymerization from taking place. Little or no polymer can be formed in wool from the base of a staple, whereas polymer can be synthesized quite readily in the tip where oxidation products are more likely to be present. If, however, wool from the base is pretreated with a mild oxidizing agent, polymerization will then occur as readily as in the tip. The polymer is apparently chemically bound to the wool through thio-ether linkages.

## I. INTRODUCTION

If wool is treated with a reducing agent and then immersed in an aqueous solution of a monomer, polymerization of the latter takes place within the fibre(1). This reaction probably involves reduced disulphide groupings in the wool, and to study its mechanism, experiments have now been carried out on the effect of cysteine and related compounds as polymerization catalysts *in vitro*. The first section of the present paper describes these experiments, whilst the second section records further studies on the application of this work to polymer formation in wool.

## II. POLYMERIZATION IN VITRO

It has been shown by Bacon(2) and Baxendale, Evans, and Park(3) that certain oxidation-reduction systems can act as powerful polymerization catalysts. As reduced wool would contain cysteine, it seemed desirable, therefore, to ascertain from experiments *in vitro* whether this compound could form redox systems with similar catalytic properties.

The experimental technique consisted in adding the catalysts (thiol and oxidizing agent) to 5 cc. of a 20 per cent. (by volume) aqueous solution of methacrylic acid and noting whether polymerization occurred from alterations in the viscosity of the solution. In the majority of experiments the reducing agent was cysteine (as hydrochloride) but as a point of interest, some experiments were carried out using glutathione, thiourea, and thiophenol. The oxidizing

\* Wool Textile Research Laboratory, C.S.I.R.O., Geelong.

agents were ammonium persulphate, hydrogen peroxide, and cysteic acid. All reagents were of A.R. quality excepting methacrylic acid and cysteic acid which were laboratory preparations.

Table 1 records the results obtained using the specified amounts of the different catalysts. Each catalyst has been classified according to its effectiveness, as follows:

- ++ Very effective: solution thickened appreciably within 15 minutes and usually set solid within 90 minutes.
- + Effective: alterations in viscosity slower than above.
- Ineffective: no apparent change in viscosity after 24 hours.

TABLE 1  
EFFECT OF THIOL COMPOUNDS AS CATALYSTS

Compound	Amount Added (g.)	Oxidizing Agent	Amount Added (g.)	Temperature (°C.)	Catalytic Effect
Cysteine .. ..	0.010	H <sub>2</sub> O <sub>2</sub> *	0.008	22	+
Cysteine .. ..	0.010	"	0.008	50	++
Cysteine .. ..	0.100	"	0.008	22	++
Cysteine .. ..	0.010	(NH <sub>4</sub> ) <sub>2</sub> S <sub>2</sub> O <sub>8</sub>	0.015	22	++
Cysteine .. ..	0.010	Cysteic acid	0.015	22	—
Cysteine .. ..	0.010	" "	0.015	50	—
Cysteine .. ..	0.010	Nil	—	50	—
Nil .. ..	—	(NH <sub>4</sub> ) <sub>2</sub> S <sub>2</sub> O <sub>8</sub>	0.015	50	—
Glutathione ..	0.010	H <sub>2</sub> O <sub>2</sub> *	0.008	50	++
Glutathione ..	0.003	(NH <sub>4</sub> ) <sub>2</sub> S <sub>2</sub> O <sub>8</sub>	0.013	50	+
Thiourea .. ..	0.015	(NH <sub>4</sub> ) <sub>2</sub> S <sub>2</sub> O <sub>8</sub>	0.015	50	+
Thiophenol ..	0.015	H <sub>2</sub> O <sub>2</sub> *	0.003	50	+

\* The required amount of a 10 volume solution added.

It can be seen that cysteine and the other thiol compounds form redox systems which are extremely powerful polymerization initiators.

Further information on the mechanism of the reaction was obtained by estimating thiol and disulphide sulphur in the system at various intervals during polymerization. An aqueous solution was prepared containing 0.05 g. of cysteine hydrochloride, 0.03 g. of hydrogen peroxide, and 1 g. of methacrylamide in 100 cc. Thiol and disulphide sulphur were determined on a 2 cc. aliquot by the method of Shinohara(4-6) using a Spekker absorptiometer. The solution was then allowed to stand at 22 °C. for 3 hours, samples being withdrawn for analysis as required.

From Figure 1 it is seen that the thiol content of the solution rapidly decreases during the first half hour whilst the disulphide content rises rapidly. The most significant feature of the results is that the overall amount of thiol and disulphide sulphur diminishes during the reaction. Thus, after 3 hours, the total sulphur present in solution as thiol and disulphide has dropped from



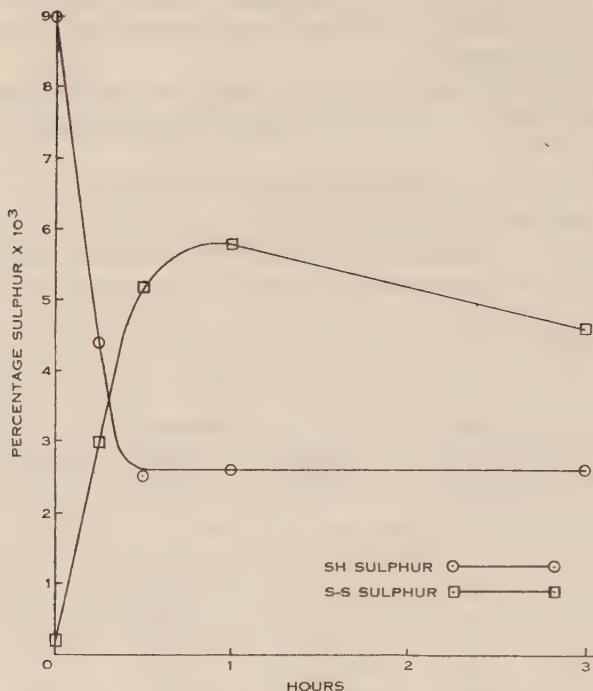


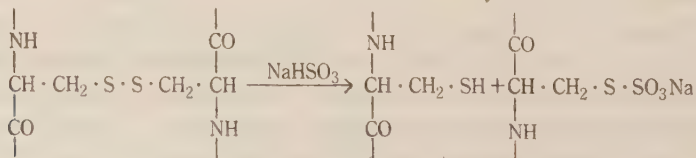
Fig. 1.—Alterations in SH and S-S sulphur during polymerization *in vitro*.

$9.25 \times 10^{-3}$  to  $7.20 \times 10^{-3}$  per cent. In other words, 22 per cent. of the original sulphur is unaccounted for either as thiol or disulphide. As in the absence of monomer no such loss is observed, it seems likely that some of the cysteine is attached to the polymer as end-groups which have acted as initiators or terminators during polymerization.

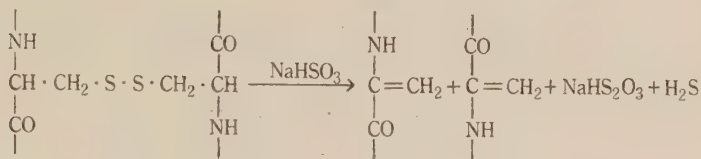
### III. POLYMERIZATION IN WOOL

From the above experiments, it appeared possible that a redox system involving cysteine was responsible for the polymerization which occurs when wool is reduced with sodium bisulphite and immersed in a solution of a monomer.

Middlebrook and Phillips(7) have divided the combined cysteine of wool into four subfractions of different reactivity with sodium bisulphite. Subfractions *A* and *B* react as follows in the cold :



the reaction being reversible for subfraction *A*. Subfraction *C* is unreactive, whilst subfraction *D* decomposes in hot bisulphite giving combined  $\alpha$ -aminoacrylic acid:



It is possible that the double bond from combined  $\alpha$ -aminoacrylic acid might initiate polymerization of ethylenic monomers in wool but, as pretreatment with *cold* bisulphite is effective, this seemed unlikely. However, to make certain, an experiment was carried out in which wool was treated with alkali which Lindley and Phillips(8) have shown converts fraction (*A* + *B*) to lanthionine and fraction (*C* + *D*) to  $\alpha$ -aminoacrylic acid.

A sample of 64's all wool flannel purified by successive extractions with light petroleum, alcohol, and water, and conditioned at 22 °C. and 65 per cent. R.H. was used in this and subsequent experiments.

A sample (2.5 g.) was treated in 1 litre of 0.1N caustic soda at 22 °C. for 6 hours and then washed overnight in running water. It was then placed in 50 cc. of 2 per cent. sodium bisulphite, raised to the boil in 5 minutes and boiled for 3 minutes. After washing overnight in running water, the sample was treated in 50 cc. of 4 per cent. (by volume) methacrylic acid for 16 hours at 22 °C. No increase in weight was recorded as a result of this treatment, although a control in which pretreatment with caustic soda was omitted showed a weight increase of 11.2 per cent. due to polymer deposition.

It is apparent, therefore, that when the cystine of fraction (*A* + *B*) is converted to lanthionine, polymerization cannot occur even though fraction (*C* + *D*) has been converted to  $\alpha$ -aminoacrylic acid. Thus it appears that cysteine formed from fraction (*A* + *B*) is responsible for the polymerization.

As shown in the previous section of this paper, a trace of oxidizing agent is necessary for cysteine to be effective as a polymerization catalyst *in vitro*. In wool, this probably occurs as an oxidation product of cystine formed during atmospheric exposure of the fleece. This is supported by the following experiments which show that little or no polymer can be formed in wool from the base of a staple, whereas considerable amounts are formed in wool from the tip of the same staple which would have been subjected to atmospheric oxidation during growth.

A sample of 70's Merino fleece wool was divided into tip and base by means of scissors. Both portions were purified by successive extractions with light petroleum, alcohol, and distilled water and conditioned at 22 °C. and 65 per cent. R.H. Samples (1 g.) from each were treated in 20 cc. of 15 per cent. sodium bisulphite for 3 hours at 22 °C., washed overnight, and placed in 50 cc.

of 4 per cent. methacrylamide for 3 hours at 50 °C. The samples were again washed overnight in running water, conditioned, and weighed. Table 2 shows the polymer contents of the wool in each series.

TABLE 2  
POLYMER CONTENTS OF TREATED WOOLS

Sample				Polymer Content (%)
Base	..	..	..	4.0
Base	..	..	..	7.5
Base	..	..	..	Nil
Tip	..	..	..	18.8
Tip	..	..	..	16.4
Tip	..	..	..	11.1
Tip	..	..	..	8.9

These results indicate that oxidation products formed in the fleece during exposure play a part in catalysing polymerization.

It seemed likely, therefore, that if wool from the base of a staple were mildly oxidized before treatment, polymerization might occur more readily. To test this, a series of experiments was next undertaken using wool from the base of staples but pretreating each sample with hydrogen peroxide before bisulphite treatment. Each sample (1 g.) was immersed in 20 cc. of hydrogen peroxide at pH 7-8 and 50 °C. and washed overnight before being treated as in the previous experiments. The conditions of treatment with hydrogen peroxide and amounts of polymer formed are given in Table 3.

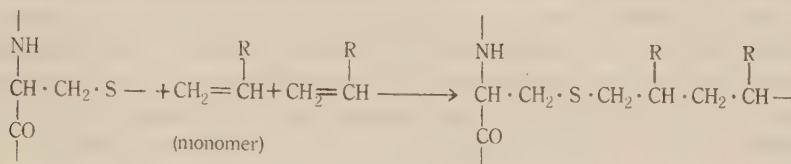
TABLE 3  
EFFECT OF PEROXIDE PRETREATMENT ON POLYMERIZATION

Strength of H <sub>2</sub> O <sub>2</sub> (%)	Time of Treatment (hr.)	Polymer Content (%)
0.60	1	16.7
0.15	1	16.2
0.03	1	1.1
0.03	2	3.0
0.03	3	2.7
0.03	1	8.4*
0.03	2	9.5*
0.03	3	11.5*

\* Washed for 3½ hours after bisulphite treatment.

It can be seen that mild pretreatment with hydrogen peroxide modified the wool from the base so that substantial quantities of polymer can now be formed. The overall result is more effective if the time of washing after bisulphite treatment is shortened to  $3\frac{1}{2}$  hours. This probably results from subfraction *A* of the cystine being rebuilt on prolonged washing. It should be mentioned, however, that if the wool is not washed at all after bisulphite treatment, polymerization is inhibited due possibly to reaction between monomer and residual bisulphite. The results shown in Table 3 are not due to residual hydrogen peroxide in the wool, as prolonged washing after pretreatment followed by immersion in bisulphite would have removed all this compound.

These experiments show that for polymerization to occur, some oxidation product must be present in the wool as well as reduced cystine from fraction (*A* + *B*). It seems unlikely that this oxidation product is cysteic acid, as it was shown earlier that a mixture of cysteic acid and cysteine does not initiate polymerization *in vitro*. For this reason, it is suggested that the active compound is probably a sulfoxide or sulphone formed by partial oxidation of some of the disulphide groupings in the wool. By reaction with reduced cystine from fraction (*A* + *B*), free radicals are probably produced which can then initiate chain polymerization as follows:



Chain termination could occur through reaction with other free radicals in the system or by chain transfer, for example, with thiol groups which are known to function in this capacity(9).

In this way, the polymer becomes chemically attached to the wool through thio-ether linkages and so a permanent union of the two is effected.

### III. REFERENCES

- (1) LIPSON, M.—*Nature* **164**: 576 (1949).
- (2) BACON, R. G. R.—*Trans. Faraday Soc.* **42**: 140 (1946).
- (3) BAXENDALE, J. H., EVANS, M. G., and PARK, G. S.—*Trans. Faraday Soc.* **42**: 155 (1946).
- (4) SHINOHARA, K.—*J. Biol. Chem.* **109**: 665 (1935).
- (5) SHINOHARA, K.—*J. Biol. Chem.* **112**: 671 (1936).
- (6) SHINOHARA, K.—*J. Biol. Chem.* **112**: 683 (1936).
- (7) MIDDLEBROOK, W. R., and PHILLIPS, H.—*Biochem. J.* **36**: 428 (1942).
- (8) LINDLEY, H., and PHILLIPS, H.—*Biochem. J.* **39**: 17 (1945).
- (9) SNYDER, H. R., STEWART, J. M., ALLEN, R. E., and DEARBORN, R. J.—*J. Amer. Chem. Soc.* **68**: 1422 (1946).



# SOME BRIDGED DERIVATIVES OF 4-PIPERIDONE

By E. F. L. J. ANET,\* G. K. HUGHES,\* DIANA MARMION,\*  
and E. RITCHIE\*

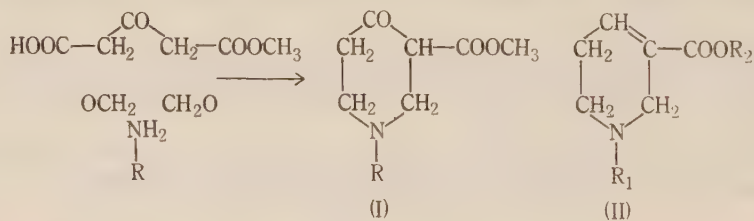
[Manuscript received December 12, 1949]

## Summary

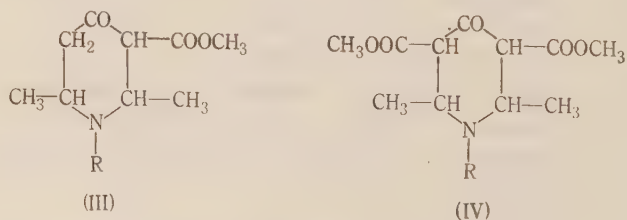
Attempts to prepare the areca nut alkaloids by the route suggested by Robinson were unsuccessful. Several polycyclic bases related to 4-piperidone have been synthesized.

## I. INTRODUCTION

According to Robinson's theory of the biogenesis of alkaloids(1) the areca nut alkaloids, guvacine (II,  $R_1=R_2=H$ ), guvacoline (II,  $R_1=H$ ,  $R_2=CH_3$ ), arecaidine (II,  $R_1=CH_3$ ,  $R_2=H$ ), and arecoline (II,  $R_1=R_2=CH_3$ ) are formed from the intermediate I ( $R=H$  or  $CH_3$ ) by reduction and dehydration and also hydrolysis in the cases of guvacine and arecaidine. Since these steps have been accomplished *in vitro*(2) and are presumably possible *in vivo*, the formation of I ( $R=H$  or  $CH_3$ ) is the key process in Robinson's scheme. He suggested that it is produced by the condensation of formaldehyde and methylamine (or ammonia) with monomethyl acetonedicarboxylate by a reaction of the Petrenko-Kritschenko type(3). Later, Mannich(4) showed that salts of primary aliphatic



amines or ammonia react smoothly with acetaldehyde and monomethyl or dimethyl acetonedicarboxylate in aqueous alcoholic solution at room temperature to yield the 4-piperidones III or IV respectively, but he reported no experiments using formaldehyde.

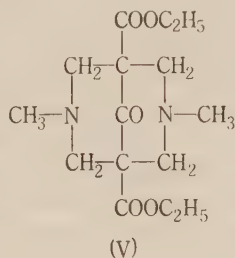


\* School of Chemistry, University of Sydney.

Preliminary attempts to obtain a 4-piperidone of type I from the reaction between formaldehyde, methylamine, and monoethyl acetonedicarboxylate (which is more readily available than the monomethyl ester) soon showed that the desired product was formed in very low yield, if at all. In the hope of determining the optimum conditions for this reaction, attention was then turned to the reaction between formaldehyde, methylamine, and diethyl acetonedicarboxylate, which was expected to be less complicated, but otherwise similar. From reactions conducted above pH 6 there was obtained much non-basic material together with smaller amounts of a dark basic oil from which no pure substance could be isolated. However, between pH 6 and pH 1.4 the basic fraction was a pale brown oil from which a substance,  $C_{15}H_{24}O_5N_2$ , was separated either by distillation or better, through its *di-hydrobromide tri-hydrate*. The yield of this salt varied according to the pH as follows:

pH	..	..	..	6.0	3.4	3.0	2.4	1.4
Yield (%)	..	..	..	0.6	3.8	7.4	6.4	3.8

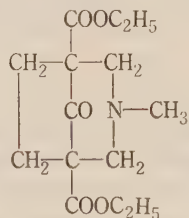
Since the substance gave a negative ferric test (absence of enolic structure) it is formulated as the bispidine, *diethyl 3,7-dimethyl-3,7-diazabicyclo[3,3,1]nonan-9-one-1,5-dicarboxylate* (V)(5). The crude basic fractions from which V was obtained always gave weakly positive ferric tests indicating that the desired diethyl 1-methyl-4-piperidone-3,5-dicarboxylate was probably present but repeated attempts to isolate it were fruitless.



In spite of this unpromising result the investigation of the reaction between formaldehyde, methylamine, and monoethyl acetonedicarboxylate (and also ethyl acetoacetate) was resumed and studied exhaustively but in no case could a pure substance be isolated.

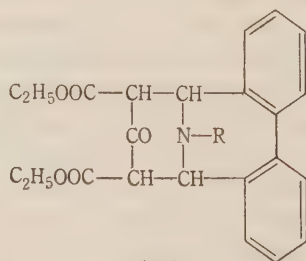
While this work was in progress there appeared a report by Herzog(6) of a lecture delivered by Schöpf, which although containing several obvious mis-statements (or misprints) does state definitely that Schöpf has likewise failed to prepare I ( $R=\text{CH}_3$ ) by the reaction under consideration. It also indicates that he has rejected it as a possible step in the biogenesis of the areca nut alkaloids and considers it more probable that they are derived from arecaine aldehyde (or its *nor*-derivative) formed by a reaction discovered by Mannich(7). Whilst a full discussion of this and other related points must be delayed until a full and authentic account of Schöpf's work and views is published, it may be pointed out now that Robinson's scheme, although lacking experimental support, is very closely related to several other reactions which have been experimentally realized.

The ease with which bispidines may be synthesized led us to examine the formation of other bridged 4-piperidones. Mannich and Schumann(8) found that dialkyl substituted acetonedicarboxylic esters react with formaldehyde and primary aliphatic amines to give derivatives of *cis*-3,5-dialkyl-4-piperidone. In a similar manner diethyl *cyclopentanone*-2,5-dicarboxylate, prepared in very low yield by the method of Mitchell and Thorpe(9), condensed readily with formaldehyde and methylamine in aqueous alcoholic solution at room temperature to yield *diethyl 3-methyl-3-azabicyclo [3,2,1] octan-8-one-1,5-dicarboxylate* (VI). This substance, which gave a negative ferric test, was characterized by its *picrate*. Construction of scale models shows that the ring system must have the *cis* configuration.



(VI)

Mannich and Veit(10) have prepared a 4-piperidone bridged across the 2 and 6 carbon atoms by an ethylene residue (dimethyl 8-methyl-8-azabicyclo [3,2,1] octan-3-one-2,4-dicarboxylate) by condensing succinaldehyde with methylamine and dimethyl acetonedicarboxylate. Similarly it was found that 4-piperidones bridged across these positions by an aromatic residue could be prepared. When an alcoholic solution of diphenyl-2,2'-dialdehyde and diethyl acetonedicarboxylate was treated at room temperature with aqueous methylamine, the nearly pure product, *diethyl 10-methyl-2,3,4,5-dibenzo-10-azabicyclo [4,3,1]-deca-2,4-dien-8-one-7,9-dicarboxylate* (VII, R=CH<sub>3</sub>) crystallized out almost immediately and eventually a nearly quantitative yield was obtained. The *10-ethyl* (VII, R=C<sub>2</sub>H<sub>5</sub>), *10-allyl* (VII, R=C<sub>3</sub>H<sub>5</sub>), *10-(2'-hydroxyethyl)* (VII, R=CH<sub>2</sub>CH<sub>2</sub>OH), *10-benzyl* (VII, R=CH<sub>2</sub>C<sub>10</sub>H<sub>5</sub>), and *ethylene-bis* (from ethylene diamine) derivatives were equally readily prepared. Indeed so facile is the reaction that it could be used as a test for primary aliphatic amines. All of these substances gave positive ferric tests. Scale models again show that the ring system must have the *cis* configuration.



(VII)

The diphenyl-2,2'-dialdehyde was prepared by a slight modification of the method of Rapson and Shuttleworth(11) from *o*-iodobenzaldehyde which was readily available from *o*-iodobenzyl bromide through the Sommelet reaction. The latter substance was obtained from *o*-iodotoluene by bromination with *N*-bromosuccinimide or better according to the excellent procedure of Cook, Dickson, and Loudon(12).

## II. EXPERIMENTAL

### (a) *Diethyl 3,7-Dimethyl-3,7-Diazabicyclo [3,3,1] Nonan-9-One-1,5-Dicarboxylate (V)*

Aqueous formaldehyde (20 ml. of 30% ; 2 mol.) was added to a solution of methylamine hydrochloride (6.75 g. ; 1 mol.) and diethyl acetonedicarboxylate (20.2 g. ; 1 mol.) in the aqueous alcoholic buffer solution (750 ml.). The initially clear colourless mixture soon became yellow and an oil began to separate. After standing for 3 days at room temperature alcohol was distilled off under reduced pressure (bath temperature <50 °C.) and the residue, after being acidified to Congo red if necessary, extracted several times with ether to remove non-basic material. Then the aqueous layer was saturated with potassium carbonate and extracted with ether. Evaporation of the ether (bath temperature <50 °C.) left a pale brown oil which gave a weak purple-red ferric test. The oil was dissolved in absolute alcohol (5 ml.), cooled in ice, and an equal bulk of cold concentrated hydrobromic acid added. After standing for several hours at 0 °C., the colourless crystalline precipitate was collected and washed with a little cold aqueous alcoholic hydrobromic acid. The yields obtained have been set out in the introduction. The substance crystallized from acetone containing a little absolute alcohol in small colourless needles, which sintered at 124 °C. and decomposed at 146 °C.

Found : C, 33.9 ; H, 6.0 ; N, 5.0 ; Br, 30.2%, in material dried for 12 hours in an evacuated desiccator over concentrated sulphuric acid.

Calculated for  $C_{15}H_{24}O_5N_2 \cdot 2HBr \cdot 3H_2O$  : C, 34.1 ; H, 6.1 ; N, 5.1 ; Br, 30.3%.

The base, regenerated from its hydrobromide, crystallized from light petroleum (b.p. <40 °C.) in colourless prismatic needles, m.p. 64 °C.

Found : C, 57.6 ; H, 7.8 ; N, 9.3%.

Calculated for  $C_{15}H_{24}O_5N_2$  : C, 57.7 ; H, 7.7 ; N, 9.3%.

The substance may be isolated from the brown oil above by distillation under reduced pressure but in diminished yield. It had b.p. 172–175 °C./1 mm. and solidified immediately.

### (b) *Diethyl 3-Methyl-3-Azabicyclo [3,2,1] Octan-8-One-1,5-Dicarboxylate (VI)*

When a solution of diethyl cyclopentanone-2,5-dicarboxylate (7.5 g. ; 1 mol.) in alcohol (10 ml.) and aqueous formaldehyde (6.7 ml. of 30% ; 2 mol.) was treated with aqueous methylamine (3.3 ml. of 30% ; 1 mol.) the temperature rose to about 40 °C. After standing for 3 days at room temperature, alcohol was removed by distillation under reduced pressure, the residue acidified to Congo red and non-basic substances extracted with ether. The base liberated from the aqueous layer by adding excess potassium carbonate was eventually isolated by distillation. It formed a colourless oil (4.7 g. ; 50%) of b.p., 185–187 °C./16 mm.

The *picrate* crystallized from alcohol in yellow prismatic needles, m.p. 130–131 °C. with decomposition.

Found : C, 46.7 ; H, 4.7 ; N, 10.9%.

Calculated for  $C_{20}H_{24}O_{12}N_4$  : C, 46.9 ; H, 4.7 ; N, 10.9%.

### (c) *o-Iodobenzaldehyde*

A gently refluxing mixture of *o*-iodotoluene (218 g. ; 1 mol.), carbon tetrachloride (800 ml.), and water (200 ml.) containing a little iodine was irradiated by three 75-watt lamps and treated during 2 hours with a solution of bromine (160 g. ; 1 mol.) in carbon tetrachloride (300 ml.). After refluxing for 1 hour longer, the solvents were distilled off from the water-bath as far as



possible, the organic layer separated from the water, diluted with chloroform (150 ml.), and then heated in a boiling water-bath under a water-pump vacuum until all the chloroform and water had been removed.

After cooling, the residue was added to a solution of hexamine (200 g.) in chloroform (2,000 ml.) and the mixture allowed to stand overnight. The addition compound was collected, washed with a little chloroform, air-dried for about 1 hour, and then refluxed with glacial acetic acid (300 ml.) and water (300 ml.) for  $1\frac{1}{2}$  hours. The mixture was cooled, diluted, and extracted several times with ether. The ethereal extracts were washed with sodium carbonate solution, with water, and then dried and distilled in a stream of nitrogen. The product (81 g.) was a pale yellow oil, b.p. 143–148 °C./28 mm. which solidified completely to a colourless solid, m.p. 35 °C. (lit. 37 °C.).

By working up the chloroform filtrates *o*-iodotoluene (76 g.) was recovered. Allowing for this the overall yield of *o*-iodobenzaldehyde was 54%.

By washing the product from the bromination with dilute sodium carbonate solution and water and then fractionating, *o*-iodotoluene (85 g.) was recovered and *o*-iodobenzylbromide (145 g.; 80%), b.p. 114–120 °C./30 mm., m.p. 50–51 °C. (lit. 55 °C.) isolated. No advantage was gained by isolating this intermediate and its powerful lachrymatory properties made it unpleasant to handle.

(d) *Diphenyl-2,2'-Dialdehyde*

When *o*-iodobenzaldehyde and copper powder were slowly heated to 200 °C. according to the method of Rapson and Shuttleworth(11) the reaction was vigorous and uncontrollable. However, if the copper powder was added during  $\frac{1}{2}$  hour to *o*-iodobenzaldehyde maintained at 200–210 °C. the reaction proceeded smoothly and the product was obtained in 55–60%.

(e) *Diethyl 10-Methyl-2,3,4,5-Dibenzo-10-Azabicyclo [4,3,1]-Deca-2,4-Dien-8-One-7,9-Dicarboxylate (VII, R=CH<sub>3</sub>)*

A solution of diphenyl-2,2'-dialdehyde (4.2 g.; 1 mol.) and diethyl acetonedicarboxylate (4 g.; 1 mol.) in alcohol (20 ml.) was cooled to 10 °C. and treated with aqueous methylamine (2 ml. of 30%; 1 mol.). The solution became faintly yellow, the temperature rose to 25–30 °C. and in a few minutes the product began to crystallize out. After standing overnight it was collected and recrystallized from alcohol from which it separated as colourless prisms, m.p. 118 °C. (7.4 g.; 90%).

Found: C, 70.5; H, 6.0%.

Calculated for C<sub>24</sub>H<sub>25</sub>O<sub>5</sub>N: C, 70.8; H, 6.2%.

It was easily soluble in benzene but much less so in alcohol. It was soluble in hot dilute acid.

The following substances were prepared in the same manner.

VII (R=C<sub>2</sub>H<sub>5</sub>), crystallized from alcohol in colourless prisms (90%), m.p. 105 °C.

Found: C, 71.0; H, 6.5%.

Calculated for C<sub>25</sub>H<sub>27</sub>O<sub>5</sub>N: C, 71.3; H, 6.5%.

VII (R=C<sub>3</sub>H<sub>7</sub>), crystallized from alcohol in colourless prisms (86%), m.p. 106 °C.

Found: C, 71.8; H, 6.4%.

Calculated for C<sub>26</sub>H<sub>27</sub>O<sub>5</sub>N: C, 72.1; H, 6.3%.

VII (R=CH<sub>2</sub>CH<sub>2</sub>OH), crystallized from alcohol in colourless needles (85%), m.p. 144 °C.

Found: C, 68.6; H, 6.2%.

Calculated for C<sub>25</sub>H<sub>27</sub>O<sub>6</sub>N: C, 68.6; H, 6.2%.

VII (R=CH<sub>2</sub>C<sub>6</sub>H<sub>5</sub>), crystallized from alcohol-benzene in colourless needles (92%), m.p. 148 °C.

Found: C, 74.4; H, 6.1%.

Calculated for C<sub>30</sub>H<sub>29</sub>O<sub>5</sub>N: C, 74.5; H, 6.0%.

The *ethylene-bis* derivative crystallized from alcohol-benzene in colourless prisms (82%) softening at about 205 °C. and decomposing at about 212 °C.

Found: C, 70.6; H, 5.9%.

Calculated for  $C_{48}H_{48}O_{10}N_2$ : C, 70.9; H, 5.9%.

All these substances gave deep wine red colours with ferric chloride in alcoholic solution.

### III. ACKNOWLEDGMENTS

The authors are indebted to Miss J. Fildes and Miss B. Naylor for the analyses, and to the Commonwealth Research Grants Committee of the University of Sydney for the award of a scholarship to one of them (E.F.L.J.A.).

### IV. REFERENCES

- (1) ROBINSON, R.—*J. Chem. Soc.* **111**: 876 (1917).
- (2) DANKOVA, T. F., SIDOROVA, E. A., and PREOBRACHENSKI, N. A.—*Chem. Abstr.* **37**: 381 (1943).
- (3) PETRENKO-KRITSCHENKO, P. I., and ZONEFF, N.—*Ber. dtsh. chem. Ges.* **39**: 1358 (1906).
- (4) MANNICH, C.—*Arch. Pharm. Berl.* **272**: 323 (1934).
- (5) MANNICH, C., and MOHS, P.—*Ber. dtsh. chem. Ges.* **63**: 608 (1930).
- (6) HERZOG, E.—*Chimia* **2**: 206 (1948).
- (7) MANNICH, C.—*Ber. dtsh. chem. Ges.* **75**: 1480 (1942).
- (8) MANNICH, C., and SCHUMANN, P.—*Ber. dtsh. chem. Ges.* **69**: 2299 (1936).
- (9) MITCHELL, A. D., and THORPE, J. F.—*J. Chem. Soc.* **97**: 997 (1910).
- (10) MANNICH, C., and VEIT, F.—*Ber. dtsh. chem. Ges.* **68**: 506 (1935).
- (11) RAPSON, W. S., and SHUTTLEWORTH, R. G.—*J. Chem. Soc.* **1941**: 487 (1941).
- (12) COOK, J. W., DICKSON, G. T., and LOUDON, J. D.—*J. Chem. Soc.* **1947**: 746 (1947).

# A SYNTHESIS OF ISOPELLETIERINE AND METHYLISOPELLETIERINE\*

By E. F. L. J. ANET,† G. K. HUGHES,† and E. RITCHIE†

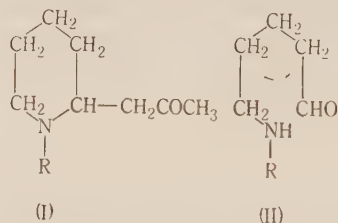
[Manuscript received December 12, 1949]

## Summary

The syntheses of *isopelletierine* and *methylisopelletierine* from 5-aminopentanal and 5-methylaminopentanal respectively, are described.

## I. INTRODUCTION

In Robinson's(1) theory of the biogenesis of alkaloids *methylisopelletierine* (I, R=CH<sub>3</sub>) is regarded as the precursor of *isopelletierine* (I, R=H), *coniine*, *methylconiine*, *conhydrine*, and  $\gamma$ -*coniceine*, giving rise to them by demethylation, reduction, dehydration, and methylation reactions. Interest in the biosynthesis of this group therefore centres primarily on the formation of *methylisopelletierine*, which Robinson suggests results from the condensation of 5-methylaminopentanal (II, R=CH<sub>3</sub>) or its ring tautomer with acetone-dicarboxylic acid. Syntheses of *isopelletierine* and *methylisopelletierine* by conventional methods have been recorded(2).



The synthesis of II (R=CH<sub>3</sub>) was not readily effected and several possible routes were explored before success was achieved. A possible route through 2, 3, 4, 5-tetrahydropyridine was unattractive and was not attempted because it was by no means certain that it could be methylated or converted to the acetal of II (R=H) and then methylated, and moreover the literature on the substance was confused. Schöpf *et al.*(3) have since clarified considerably its chemistry. As the original paper by Schöpf *et al.* on tetrahydropyridine was not available to us our first attempt to synthesize II was by reduction of 3-bromopropyl-cyanide by Stephen's(4) method to 4-bromobutanal, which should be convertible to II (R=CH<sub>3</sub>) by the reactions used for the preparation of its lower homologue(5), but the substance was recovered unchanged. The second route

\* A short account of this work has appeared in *Nature* **164**: 501 (1949).

† School of Chemistry, University of Sydney.

was based on the claim(6) that acetaldehyde can be cyanoethylated to a mixture of 4-cyanobutanol and 2,2'-cyanoethyl-4-cyanobutanol in a total yield of 40-50 per cent. It was hoped to transform the former substance to its acetal, then reduce and methylate to the acetal of II ( $R=CH_3$ ) but several trials gave the desired product in very impure form in yields of less than 5 per cent. Next it was proposed to brominate 1,1-diethoxy-2-butene, for which an improved preparation was devised, with *N*-bromosuccinimide to 4-bromo-1,1-diethoxy-2-butene, convert this to the related cyanide, reduce fully, and methylate to the acetal of II ( $R=CH_3$ ). Ziegler *et al.*(7) reported that 1,1-diethoxy-2-butene reacted with *N*-bromosuccinimide but did not describe the products. Later, Flaig(8) who obtained 4-bromo-1,1-diethoxy-2-butene by another method, found that it was a very unstable substance, but since no reason for this instability was apparent and since his original paper was not available to us, the bromination was attempted. Under a variety of conditions succinimide was recovered quantitatively but fractional distillation at 1 mm. of the remainder of the reaction mixture failed to yield the desired product (see 9).

Eventually a synthesis was devised starting from tetrahydrofurfuryl alcohol which was converted to 4-chlorobutanol by the method of Paul(10). Treatment of this with anhydrous ethanol containing dry hydrogen chloride or better with ethyl orthoformate also, gave 4-chloro-1,1-diethoxybutane (III) which readily yielded 4-cyano-1,1-diethoxybutane (IV) (see 11). Reduction of IV by sodium and alcohol afforded 5-amino-1,1-diethoxypentane (V) which was then methylated through its *benzylidene derivative* (VI) to 5-methylamino-1,1-diethoxypentane (VII) by the method used for its lower homologue(5).

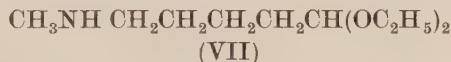
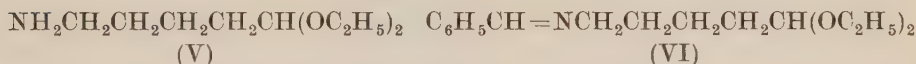
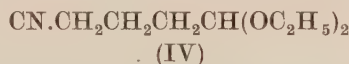


TABLE 1

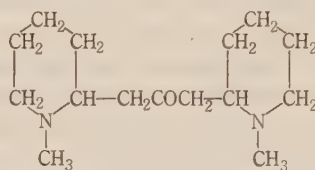
	Synthetic Product	Methylisopelletierine
Base (b.p.) .. ..	100 °C./15 mm.	95 °C./12 mm.*
Picrate (m.p.) .. ..	155 °C.	156 °C.*
Hydrobromide (m.p.) ..	151 °C.	150-151 °C.*

\* Meisenheimer and Mahler(12). These are the constants of a synthetic base, rigorously purified, whose identity was fully established by direct comparison with an authentic specimen and they are undoubtedly the best in the literature. The values for the natural alkaloid given by Hess (see (12) and numerous other papers) are rather variable.



When a faintly acid solution of VII was warmed for a short time, then cooled, neutralized, and added to a dilute solution of acetoacetic acid in a buffer of pH 7, a base,  $C_9H_{17}ON$ , was produced in high yield after 24 hours at room temperature. The comparisons of melting points and boiling points listed in Table 1 identify this synthetic base as methylisopelletierine.

In the same manner, condensation of VII with acetonedicarboxylic acid produced 1,3-bis-2'-(1'-methylpiperidyl)propanone (VIII) an analogue of cuscohygrine, which was characterized by its *dipicrate*.



(VIII)

In the reaction between V and acetoacetic acid the yield of the product, a base,  $C_8H_{15}ON$ , depended markedly on the pH. The results are summarized :

pH	..	..	7	9	11	13
Yield (%)	..	..	2	15	54	58

The base was identified as isopelletierine by the comparisons of melting points and boiling points listed in Table 2.

TABLE 2

	Synthetic Product	<i>iso</i> Pelletierine
Base (b.p.) .. ..	94 °C./15 mm.	86 °C./10 mm.*
Picrate (m.p.) .. ..	148 °C.	147-8 °C.*
Hydrobromide (m.p.) ..	135 °C.	137 °C.*
Hydrochloride (m.p.) ..	145 °C.	144-5 °C.*

\* Meisenheimer and Mahler(12) (see footnote to Table 1).

The success of this synthesis led us to re-examine the analogous reaction between acetoacetic acid and 4-aminobutanal, which according to Schöpf(13) does not yield *nor*hygrine. However, experiments conducted over the range pH 7 to pH 13 confirmed his results.

In a report by Herzog(14) of a lecture delivered by Schöpf it is stated that he also has synthesized *isopelletierine* by condensing 2,3,4,5-tetrahydropyridine with acetoacetic acid at pH 11, but no mention is made of a synthesis of methylisopelletierine.

## II. EXPERIMENTAL

(a) *1,1-Diethoxy-2-Butene*

A solution of 2-butenal (35 g.; 1 mol.) and ethyl orthoformate (90 g.; 1.2 mol.) in anhydrous ethanol (70 g.; 3 mol.) previously saturated with ammonium chloride at its boiling point, was refluxed for 10 minutes, cooled, diluted with ether,\* and then shaken with water. The product (54 g.; 75%) isolated from the ether layer by fractional distillation had b.p. 146–148 °C. (lit. 147–148 °C.).

(b) *4-Chloro-1,1-Diethoxybutane (III)*

(i) A solution of 4-chlorobutanal\* (54 g.; 1 mol.) in anhydrous ethanol (230 g.; 10 mol.) containing dry hydrogen chloride (2.3 g.) was allowed to stand at room temperature for 5 days, then neutralized to phenolphthalein with alcoholic sodium ethoxide, filtered, and fractionally distilled, finally under reduced pressure. The product, a colourless mobile oil (13.5 g.; 15%) had b.p. 95 °C./25 mm.

(ii) A solution of 4-chlorobutanal (107.5 g.; 1 mol.) and ethyl orthoformate (148 g.; 1 mol.) in anhydrous ethanol (50 g.) containing dry hydrogen chloride (0.5 g.) was allowed to stand at room temperature for 5 days, then worked up as above. The yield (81 g.; 45%) was still poor.

Found: C, 53.9; H, 9.6%.

Calculated for  $C_8H_{17}O_2Cl$ : C, 53.2; H, 9.4%.

(c) *4-Cyano-1,1-Diethoxybutane (IV)*

4-Chloro-1,1-diethoxybutane (181 g.; 1 mol.), potassium cyanide (130 g.; 2 mol.), potassium iodide (2 g.), and ethylene glycol (1,000 ml.) were stirred at 100 °C. for 4 hours, then poured into water (2,000 ml.) and extracted 8 times with ether. The combined ethereal extracts after washing with a little water and drying over sodium sulphate were distilled. The product (128–135 g.; 75–80%) had b.p. 120 °C./9 mm.

Found: N, 8.0%.

Calculated for  $C_9H_{17}O_2N$ : N, 8.2%.

(d) *5-Amino-1,1-Diethoxypentane (V)*

Sodium (26 g.) was added in one lot to a solution of 4-cyano-1,1-diethoxybutane (30 g.) in anhydrous ethanol (400 ml.). When the reaction had ceased, water (80 ml.) was added, and the mixture evaporated as far as possible on the water-bath under reduced pressure. Then water (150 ml.) was added and the amine extracted several times with ether. Fractionation under reduced pressure gave a colourless oil (24 g.; 80%), b.p. 73 °C./1.5 mm.

Found: N, 7.8%.

Calculated for  $C_9H_{21}O_2N$ : N, 8.0%.

(e) *Benzylidene-5-Amino-1,1-Diethoxypentane (VI)*

5-Amino-1,1-diethoxypentane (17.5 g.; 1 mol.), freshly distilled benzaldehyde (11.7 g.; 1.1 mol.), and anhydrous potassium carbonate (2.8 g.; 0.2 mol.) were allowed to stand overnight. After adding ether, the mixture was filtered and the filtrate distilled. The product, a colourless viscous oil, b.p. 143 °C./0.3 mm., was obtained in nearly quantitative yield.

Found: C, 72.8; H, 9.5; N, 5.4%.

Calculated for  $C_{16}H_{25}O_2N$ : C, 73.0; H, 9.7; N, 5.3%.

(f) *5-Methylamino-1,1-Diethoxypentane (VII)*

Pure dry methyl iodide (31.2 g.; 2 mol.) was added to VI (29 g.; 1 mol.) and the mixture allowed to stand at room temperature for 3 days. The quaternary salt which formed a dark

\* Further work has shown that this is a mixture containing non-aldehydic material, which may account for the low yield of the acetal. The matter is being further investigated.

viscous lower layer was washed thoroughly with dry ether and then shaken with a small volume of water. After 2 minutes the separated benzaldehyde was removed with ether, the aqueous layer made just alkaline with dilute sodium hydroxide solution, and extracted once with ether (E1). The aqueous liquid was then made strongly alkaline by the addition of excess concentrated sodium hydroxide solution and extracted 4 times with ether (E2). Distillation of E2 gave eventually VII (7.1 g.; 35%) as a colourless oil, b.p. 68 °C./1 mm.

Found: C, 62.9; H, 12.4; N, 7.6%.

Calculated for  $C_{10}H_{23}O_2N$ : C, 63.4; H, 12.3; N, 7.4%.

From E1 there was obtained a thick colourless oil (8.2 g.), b.p. 190 °C./1 mm. or 60 °C./10<sup>-3</sup> mm.

Found: C, 64.0; H, 11.7; N, 7.5%.

(g) *Methylisopelletierine* (I,  $R=CH_3$ )

A solution of 5-methylamino-1,1-diethoxypentane (5.7 g.; 1 mol.) in 1% hydrochloric acid, just acid to Congo red, was heated on the water-bath for 8 minutes, cooled, quickly neutralized with dilute potassium hydroxide solution, and immediately added to a neutralized solution of acetoacetic acid prepared from ethyl acetoacetate (7.8 g.; 2 mol.) and phosphate buffer (150 ml. of M/10) of pH 7. The volume was made up to 450 ml. and the mixture allowed to stand at room temperature for 24 hours. After saturating with potassium carbonate the mixture was extracted with ether 8 times. The base (3.45 g.; 73%) isolated by distillation had b.p. 100 °C./15 mm. It was analysed in the form of the derivatives described below.

The *hydrobromide* crystallized from acetone in long colourless needles, m.p. 151 °C.

Found: C, 45.4; H, 7.7; N, 5.8; Br, 34.1%.

Calculated for  $C_8H_{18}ONBr$ : C, 45.8; H, 7.7; N, 5.9; Br, 33.9%.

The *picate* crystallized from acetone in yellow needles, m.p. 155 °C.

Found: N, 14.4%.

Calculated for  $C_{15}H_{20}O_8N_4$ : N, 14.6%.

(h) *1,3-Bis-2'(1'-Methylpiperidyl)propanone* (VIII)

Condensation of VII (3.8 g.; 2 mol.) with acetonedicarboxylic acid (1.46 g.; 1 mol.) at pH 7 as above gave eventually a viscous colourless oil (1.25 g.; 50%), b.p. 142 °C./1 mm.

The dipicate crystallized from acetone in yellow prisms, m.p. 203–204 °C. (with decomp.).

Found: N, 15.7%.

Calculated for  $C_{27}H_{34}O_{15}N_8$ : N, 15.8%.

(i) *isopelletierine* (I,  $R=H$ )

The acetal (V) (5.25 g.; 1 mol.) was hydrolysed to the corresponding aldehyde in the usual way and this reacted with acetoacetic acid, prepared from ethyl acetoacetate (7.8 g.; 2 mol.), at room temperature for 48 hours in buffered solution (450 ml.). The yields are given in the introduction. The reaction at pH 13 was carried out in N/10 sodium hydroxide. In working up this reaction, the mixture was acidified to Congo red with hydrochloric acid, and boiled for 3 hours before saturating with potassium carbonate and extracting with ether. This procedure was adopted to effect decarboxylation of the intermediate substituted acetoacetic acid (see 15). The base which had b.p. 94 °C./15 mm. was analysed in the form of the derivatives described below.

The *hydrobromide* crystallized from acetone in colourless needles, m.p. 135 °C.

Found: C, 43.4; H, 7.2; N, 6.3; Br, 36.3%.

Calculated for  $C_8H_{16}ONBr$ : C, 43.2; H, 7.3; N, 6.3; Br, 36.0%.

The *hydrochloride*, colourless needles from acetone had m.p. 145 °C.

Found: N, 8.0%.

Calculated for  $C_8H_{16}ONCl$ : N, 7.9%.

The *picrate* separated from acetone in yellow needles, m.p. 148 °C.

Found: N, 15.1%.

Calculated for  $C_{14}H_{18}O_8N_4$ : N, 15.1%.

The *picrolonate* crystallized from ethanol in yellow needles, m.p. 180 °C.

Found: N, 17.1%.

Calculated for  $C_{18}H_{23}O_6N_5$ : N, 17.3%.

### III. ACKNOWLEDGMENTS

The authors are grateful to Miss J. Fildes, and Miss B. Naylor, for the analyses; to the Commonwealth Research Grant Committee of the University of Sydney for a scholarship awarded one of them (E.F.L.J.A.); and to The Quaker Oats Company through their distributors, Swift and Company Pty. Ltd., for a generous gift of tetrahydrofurfuryl alcohol.

### IV. REFERENCES

- (1) ROBINSON, R.—*J. Chem. Soc.* **111**: 876 (1917).
- (2) HENRY, T. A.—“The Plant Alkaloids.” 4th Ed. (Churchill: London, 1949).
- (3) SCHÖPF, C., KOMZAK, A., BRAUN, F., and JACOBI, E.—*Liebigs Ann.* **559**: 1 (1948).
- (4) STEPHEN, H.—*J. Chem. Soc.* **127**: 1874 (1925).
- (5) ANET, E. F. L. J., HUGHES, G. K., and RITCHIE, E.—*Aust. J. Sci. Res. A* **3**: (1950).
- (6) DU PONT DE NEMOURS & Co.—Brit. Pat. 576,427 (April 3, 1946).
- (7) ZIEGLER, K., SPAETH, A., SCHAAF, E., SCHUMANN, W., and WINKELMANN, E.—*Liebigs Ann.* **551**: 80 (1942).
- (8) FLAIG, W.—*Chem. Abstr.* **41**: 6189 (1947).
- (9) SCHMID, H., and GROB, E.—*Helv. Chim. Acta* **32**: 77 (1949).
- (10) PAUL, R.—*C.R. Acad. Sci. Paris.* **211**: 645 (1940); **215**: 303 (1942).
- (11) LA FORGE, F. B., GREEN, N., and GERSDORFF, W. A.—*J. Amer. Chem. Soc.* **70**: 3707 (1948).
- (12) MEISENHEIMER, J., and MAHLER, E.—*Liebigs Ann.* **462**: 301 (1928).
- (13) SCHÖPF, C.—*Z. angew. Chem.* **50**: 797 (1937).
- (14) HERZOG, E.—*Chimia.* **2**: 206 (1948).
- (15) SCHÖPF, C., and LEHMANN, G.—*Liebigs Ann.* **518**: 1 (1935).



# THE CHEMICAL CONSTITUENTS OF AUSTRALIAN *XANTHOXYLUM* SPECIES

## I. SUBEROSIN, 6-( $\gamma\gamma$ -DIMETHYLALLYL)-7-METHOXYCOUMARIN

By JEAN EWING,\* G. K. HUGHES,\* and E. RITCHIE\*

[Manuscript received March 10, 1950]

### Summary

*Suberosin*, a new coumarin, has been isolated from the bark of *Xanthoxylum suberosum* and shown to be 6-( $\gamma\gamma$ -dimethylallyl)-7-methoxycoumarin.

### I. INTRODUCTION

The genus *Xanthoxylum* is placed by Engler and Prantl(1) in subtribe Evodiinae, tribe Xanthoxyleae, family Rutaceae. It is a large genus and is widely spread over the tropical and subtropical regions of the world. The five Australian species belong to the subsection Blackburnia of the genus. Except for the vine *X. torvum* F. Muell. they are small or medium sized trees and are endemic to Australia. They occur in dry "scrubs" or rain-forests in northern and north-eastern Australia, while one, *X. brachyacanthum* F. Muell., extends from north Queensland into northern New South Wales.

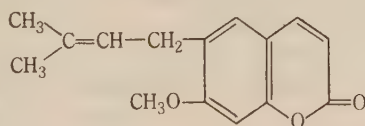
*Xanthoxylum suberosum* C. T. White is a small tree of irregular growth, which averages about 10 feet in height and is confined to the rain-forests of the Atherton Tableland. During an examination of its bark for alkaloids† it was observed that when a concentrated methanolic extract was allowed to stand in the refrigerator a colourless crystalline solid gradually separated. This substance, m.p. 87.5 °C., now named *suberosin*, analysed for  $C_{15}H_{16}O_3$  and contained one methoxyl group, but tests for the methylenedioxy group were negative. It was insoluble in water, dilute mineral acids, and aqueous sodium carbonate but slowly dissolved in boiling dilute sodium hydroxide to a yellow solution from which it was precipitated unchanged by the addition of acid. It readily formed a *dibromide* and catalytic hydrogenation furnished *tetrahydro-suberosin*. These reactions suggested that suberosin was a methoxycoumarin with a  $C_5H_9$  side chain, which on general grounds was probably  $\gamma\gamma$ -dimethylallyl. The possibility that it was identical with osthol, m.p. 83–84 °C.(2) was eliminated by direct comparison (mixed m.p.) with an authentic specimen.

The presence of a coumarin ring was established by methylation with dimethyl sulphate and alkali when there was obtained *methylsuberosinic acid*,  $C_{16}H_{20}O_4$ , which on catalytic reduction yielded *tetrahydromethylsuberosinic acid*,  $C_{16}H_{24}O_4$ . Oxidation of suberosin by chromic anhydride in hot acetic acid

\* School of Chemistry, University of Sydney.

† A report on these alkaloids will be made later.

solution gave acetone, an acid, and a small amount of the corresponding aldehyde. The acid  $C_{11}H_8O_5$  was identified by direct comparison of it and its methyl ester with authentic synthetic specimens, as 7-methoxycoumarin-6-carboxylic acid(3). When the oxidation was carried out at room temperature the aldehyde,  $C_{11}H_8O_4$ , m.p. 252–253 °C. (lit. 252–254 °C.), characterized by its oxime, m.p. 228 °C. (lit. 228 °C.) was the chief product. Suberosin thus behaves unlike osthol (8-( $\gamma\gamma$ -dimethylallyl)-7-methoxycoumarin) which yields acetone and ostholic acid,  $C_{12}H_{10}O_5$ , on oxidation(4) but like ostruthin methyl ether (6-geranyl-7-methoxycoumarin) in which a similar loss of a carbon atom occurs yielding 7-methoxycoumarin-6-aldehyde oxidizable further to 7-methoxycoumarin-6-carboxylic acid(4, 5). It follows then that suberosin is 6-( $\gamma\gamma$ -dimethylallyl)-7-methoxycoumarin (I).



Additional evidence supporting this conclusion was obtained from the permanganate oxidation of methylsuberosinic acid when 4,6-dimethoxyisophthalic acid was formed and from the alkali fusion of suberosin which yielded resorcinol and  $\beta$ -resorecylic acid.

## II. EXPERIMENTAL

### (a) Isolation of Suberosin

The dried milled bark was exhausted with methanol at room temperature and the dark extract concentrated under reduced pressure to a thin syrup. On standing in the refrigerator, a crystalline substance gradually separated. After several days it was collected and washed, firstly, with a little cold methanol and then with light petroleum. The crude product could be purified by repeated recrystallization from methanol or better by passing a dilute (about 1%) solution in benzene through a column of alumina when nearly all of the coloured impurities were preferentially adsorbed. The pure substance separated from methanol in colourless prisms, m.p. 87.5 °C. By working up mother liquors a total yield of about 1% was obtained.

Found: C, 73.5, 72.9; H, 6.9, 6.8;  $CH_3O$ , 12.8, 12.9%.

Calculated for  $C_{15}H_{16}O_3$ : C, 73.7; H, 6.6;  $1CH_3O$ , 12.7%.

It was sparingly soluble in light petroleum but fairly easily soluble in the usual organic solvents. In concentrated sulphuric acid it formed a bright orange solution with a purple fluorescence.

### (b) Suberosin Dibromide

A solution of suberosin (1 g.) in pure dry benzene (20 ml.) was treated with shaking with a 1% solution of bromine in benzene until a faint permanent colour was produced and then evaporated to dryness on the water-bath. The residue on recrystallization from alcohol afforded the product as colourless iridescent plates, m.p. 148 °C. in nearly quantitative yield.

Found: C, 44.1; H, 4.2; Br, 40.0%.

Calculated for  $C_{15}H_{16}O_3Br_2$ : C, 44.6; H, 4.0; Br, 39.6%.

(c) *Tetrahydrosuberosin*

Suberosin (0.5 g.) in ethanol (100 ml.) was hydrogenated at room temperature and pressure in the presence of palladized charcoal (0.25 g. of 10%) when slightly more than the theoretical amount of hydrogen was absorbed in about half an hour. The product crystallized from aqueous alcohol in colourless prisms, m.p. 48 °C.

Found: C, 72.2; H, 8.0%.

Calculated for  $C_{15}H_{20}O_3$ : C, 72.5; H, 8.1%.

(d) *Methylsuberosinic Acid*

Suberosin (2 g.) was stirred at 60 °C. with methanol (15 ml.) and aqueous sodium hydroxide (30 ml. of 20%) until it had completely dissolved to a bright yellow solution. Then dimethyl sulphate was gradually run in until the mixture was neutral. A further amount of aqueous sodium hydroxide (10 ml. of 20%), followed by dimethyl sulphate, was added and the whole process again repeated. Finally excess sodium hydroxide was added, the mixture refluxed for 1 hour, concentrated, and acidified. The precipitate on recrystallization from alcohol afforded pale yellow prisms, m.p. 134 °C. (1.8 g.).

Found: C, 69.5; H, 7.5;  $CH_3O$ , 21.8%.

Calculated for  $C_{16}H_{20}O_4$ : C, 69.5; H, 7.3;  $2CH_3O$ , 22.5%.

(e) *Tetrahydromethylsuberosinic Acid*

Reduction was carried out as under (c). The product separated from aqueous alcohol as colourless plates, m.p. 64 °C. in nearly quantitative yield.

Found: C, 68.3; H, 8.7%.

Calculated for  $C_{16}H_{24}O_4$ : C, 68.5; H, 8.6%.

(f) *Oxidation of Methylsuberosinic Acid*

A vigorously stirred solution of methylsuberosinic acid (0.5 g.) in very dilute potassium hydroxide (200 ml.) at room temperature was oxidized by aqueous potassium permanganate (280 ml. of 1%) which was added in 10 ml. lots, each portion being run in only after the previous portion had been consumed. The mixture was filtered, the manganese dioxide washed thoroughly and the combined filtrate and washings, after the addition of a little alcohol to destroy excess permanganate, concentrated under reduced pressure. Acidification followed by cooling, scratching, and standing gave a precipitate (0.1 g.) which on recrystallization from alcohol furnished 4,6-dimethoxyisophthalic acid as minute colourless needles, m.p. 266 °C. with decomposition (lit. 264 °C. with decomposition).

Found: C, 52.8; H, 4.6%.

Calculated for  $C_{10}H_{10}O_6$ : C, 53.1; H, 4.5%.

(g) *Oxidation of Suberosin*

A solution of chromic anhydride (10 g.) in water (20 ml.) was added gradually with shaking to a solution of suberosin (2 g.) in the minimum of warm glacial acetic acid in a flask fitted with a reflux condenser. A vigorous reaction occurred and the mixture almost boiled. When the reaction had subsided and the temperature had fallen to about 50 °C. a few ml. were distilled over from the mixture. From this distillate acetone 2,4-dinitrophenylhydrazones, m.p. 128 °C. alone or mixed with an authentic specimen, was readily isolated.

Found: N, 23.5%.

Calculated for  $C_9H_{10}O_4N_4$ : N, 23.5%.

The distilland on cooling and standing gradually deposited a crystalline solid, a little more of which was obtained by concentrating the mother liquor and then diluting with water. The total yield (about 1 g.) was shaken with dilute sodium carbonate (50 ml. of 5%) at room temperature and then, after standing for some hours, filtered. The residue (0.2 g.) on crystallization

from chloroform-light petroleum gave 7-methoxycoumarin-6-aldehyde as colourless needles, m.p. 252–253 °C.

Found: C, 64.2; H, 4.0%.

Calculated for  $C_{11}H_8O_4$ : C, 64.7; H, 4.0%.

The *phenylhydrazone* formed minute pale yellow needles, m.p. 183 °C. from alcohol.

Found: N, 9.6%.

Calculated for  $C_{17}H_{14}O_3N_2$ : N, 9.5%.

The oxime separated from aqueous alcohol in colourless needles, m.p. 228 °C. with decomposition.

Found: N, 6.3%.

Calculated for  $C_{11}H_9O_4N$ : N, 6.4%.

The alkaline filtrate obtained above, on acidification yielded 7-methoxycoumarin-6-carboxylic acid (0.7 g.) which separated from methanol in minute colourless needles, m.p. 264 °C., undepressed by admixture with an authentic specimen.

Found: C, 59.6; H, 3.8%.

Calculated for  $C_{11}H_8O_6$ : C, 60.0; H, 3.7%.

The methyl ester crystallized from methanol in colourless needles, m.p. 167 °C., undepressed by admixture with a synthetic specimen.

Found: C, 61.3; H, 4.2%.

Calculated for  $C_{12}H_{10}O_5$ : C, 61.5; H, 4.3%.

When the oxidation was carried out at room temperature the proportion of the acid to the aldehyde was reversed although the total yield remained approximately the same.

#### (h) *Alkali Fusion of Suberosin*

Suberosin (2 g.) was added slowly with stirring to moist potassium hydroxide (10 g.) fused at 250 °C. After a further hour's heating, the melt was cooled, dissolved in water, acidified, and exhausted with ether. The residue from the evaporation of the extract was dissolved in a little 5% sodium hydroxide which was then saturated with carbon dioxide. The mixture was thoroughly extracted with ether, the extract dried and evaporated. From the residue by distillation at 180 °C./0.1 mm., resorcinol, identified by its m.p. of 108 °C. alone or admixed with an authentic specimen, and by the fluorescein test, was obtained in slight amount.

The alkaline aqueous layer after acidification was extracted with ether and the ether removed. The residue on distillation at 180–200 °C./0.1 mm. gave an oily distillate which slowly solidified. Crystallization from water eventually gave a small amount of an acid which had m.p. 195–197 °C. with decomposition, when placed in a bath at 150 °C. The same behaviour was observed with authentic  $\beta$ -resorcylic acid and with a mixture of the two acids.

### III. ACKNOWLEDGMENTS

The authors are grateful to Mr. L. J. Webb, C.S.I.R.O., for the plant material and for botanical advice, and to Miss J. Fildes and Miss B. Naylor for the analyses.

### IV. REFERENCES

- (1) ENGLER, A., and PRANTL, K.—“Die natürlichen Pflanzenfamilien.” (Engelmann: Leipzig, 1931.)
- (2) SPÄTH, E., and PESTA, O.—*Ber. dtsh. chem. Ges.* **66**: 754 (1933).
- (3) VON BRUCHHAUSEN, F., and HOFFMANN, H.—*Ber. dtsh. chem. Ges.* **74**: 1584 (1941).
- (4) BUTENANDT, A., and MARTEN, A.—*Liebigs Ann.* **495**: 187 (1932).
- (5) SPÄTH, E., and KLÄGER, K.—*Ber. dtsh. chem. Ges.* **67**: 859 (1934).



# THE ALKALOIDS OF *PLEOGYNE CUNNINGHAMII*\*

By F. A. L. ANET,† G. K. HUGHES,† and E. RITCHIE†

[Manuscript received March 21, 1950]

## Summary

The roots of *Pleogyne cunninghamii* have been shown to contain the two previously described alkaloids *l*-bebeerine (*l*-curine) and *d*-isochondrodendrine.

## I. INTRODUCTION

*Pleogyne cunninghamii* Miers, order Berberidales, family Menispermaceae, is a vine fairly common in southern and central coastal Queensland along creek banks in cleared "scrub" areas, or in light "scrub" or rain-forest. There are no references to the chemistry or pharmacology of this species, although other local members of the family are suspected of being poisonous, and spot tests (L. J. Webb, personal communication) have indicated that most Australian species of the family contain alkaloids, particularly in the roots. This paper describes the isolation and identification of two alkaloids from the roots of this plant. No alkaloids were found in the leaves or the twigs.

The basic fraction obtained from a methanol extract of the root bark was separated by repeated treatment with ether or benzene. The soluble base (*A*) was identified as *l*-bebeerine (*l*-curine) whilst the insoluble fraction (alkaloid *B*) was found to be *d*-isochondrodendrine.

The identity of *A* with *l*-bebeerine (Table 1) was not apparent at first, owing to the fact that it could only be crystallized from benzene, m.p. 165–166 °C., and it was only after many attempts that it could be crystallized from methanol, having then m.p. 213–214 °C. Similar difficulty was encountered by Dutcher(1) who, however, was unable to crystallize a sample of benzene-crystallized *l*-bebeerine from methanol. King(3), nevertheless, found that his *l*-bebeerine always crystallized readily from methanol, and in the present work it also crystallized, even if impure, from that solvent once seeds were obtained.

Alkaloid *B*, which was identified as *d*-isochondrodendrine (Table 2), although only sparingly soluble in chloroform or methanol was found to be readily soluble in a mixture of these solvents and the alkaloid shows a definite *laevorotation* in that solvent mixture, whereas in pyridine or as an aqueous solution of a salt it is *dextrorotatory*.

Since the numerous values of the physical constant of these alkaloids given in the literature are not always in good agreement, the values now quoted are those of recent workers who characterized their alkaloids thoroughly.

\* This plant material was supplied by Mr. L. J. Webb, C.S.I.R.O., and this paper is part of the survey on the chemical constituents of the Australian flora being carried out by C.S.I.R.O. and other institutions.

† School of Chemistry, University of Sydney.

TABLE 1

Derivative	Alkaloid A	l-Bebeerine	Reference
Base ..	m.p. 165–166 °C. ( <i>ex</i> C <sub>6</sub> H <sub>6</sub> ) m.p. 213–214 °C. ( <i>ex</i> CH <sub>3</sub> OH)	m.p. 165–167 °C. ( <i>ex</i> C <sub>6</sub> H <sub>6</sub> ) m.p. 213 °C. ( <i>ex</i> CH <sub>3</sub> OH) m.p. 221–221.5 °C. ( <i>vac.</i> ) ( <i>ex</i> CH <sub>3</sub> OH)	(1) (3) (4)
	$[\alpha]_{\text{D}}^{23^\circ}$ —284° (CHCl <sub>3</sub> , <i>c</i> , 0.34)	$[\alpha]_{\text{D}}^{24^\circ}$ —190°* (CHCl <sub>3</sub> , <i>c</i> , 0.6)	(1)
	$[\alpha]_{\text{D}}^{23^\circ}$ —331° (pyridine, <i>c</i> , 0.4)	$[\alpha]_{\text{D}}^{24^\circ}$ —332° (pyridine, <i>c</i> , 0.8)	(1)
	$[\alpha]_{\text{D}}^{23^\circ}$ —286° (0.1N HCl, <i>c</i> , 0.4)	$[\alpha]_{\text{D}}^{24^\circ}$ —280° (0.1N HCl, <i>c</i> , 0.9)	(1)
Hydrochloride	m.p. 270 °C. (decomp.)	m.p. 271–272 °C.	(4)
Methiodide ..	m.p. 250 °C. (decomp.)	m.p. 249–250 °C. (decomp.)	(1)
Dimethyl ether ..	Amorphous	Amorphous	(4)
Nitrogen-free product from Hofmann degradation ..	m.p. 196 °C.	m.p. 198–199 °C.	(3)

\* Probably a misprint for —290°.

TABLE 2

Derivative	Alkaloid B	d-isoChondrodendrine	Reference
Base ..	m.p. 300–305 °C. (decomp.) $[\alpha]_{\text{D}}^{23^\circ}$ +54° (pyridine, <i>c</i> , 0.9) $[\alpha]_{\text{D}}^{23^\circ}$ —25° (1 : 1 vol. CHCl <sub>3</sub> –CH <sub>3</sub> OH, <i>c</i> , 0.38)	m.p. 316 °C. m.p. 305 °C. (decomp.) $[\alpha]_{\text{D}}^{22^\circ}$ +50° (pyridine, <i>c</i> , 0.97)	(2) (1) (1)
Sulphate ..	m.p. 285 °C. (decomp.)	m.p. 291–292 °C. (efferv.)	(2)
Dimethiodide	m.p. 285–287 °C. (decomp.)	m.p. 287 °C. (decomp.)	(2)
Dimethyl ether ..	m.p. 270–272 °C. $[\alpha]_{\text{D}}^{23^\circ}$ —15° (CHCl <sub>3</sub> , <i>c</i> , 1.0)	m.p. 269–270 °C. $[\alpha]_{\text{D}}^{24^\circ}$ —15° (CHCl <sub>3</sub> , <i>c</i> , 1.1)	(1)
Dimethyl ether dimethiodide	m.p. 310–312 °C.	m.p. 312 °C.	(2)
α - Methylmethine ..	m.p. 205–206 °C.	m.p. 206–207 °C.	(2)

## II. EXPERIMENTAL

(All melting points are uncorrected.)

(i) *Isolation*.—The ground roots (1.6 kg.) were percolated with methanol until tests for alkaloids were negative. The combined extracts were evaporated to dryness, the residue stirred with 5% hydrochloric acid (500 ml.), the resulting mixture filtered, the filtrate basified with

ammonia, and then shaken with chloroform. Separation of the layers was facilitated by filtration. The chloroform extract was found to contain all the alkaloid present and this was completely extracted by dilute sodium hydroxide solution (5%), but not by ammonia. The alkaline extract on saturation with carbon dioxide gave a precipitate which was filtered off and dried (20 g.).

(ii) *Separation of l-Bebeerine*.—The crude alkaloid was extracted in a Soxhlet apparatus with ether, when almost colourless *l*-bebeerine (7 g.) separated; extraction with benzene gave dark coloured crystalline *l*-bebeerine. Either product on repeated recrystallization from benzene had m.p. 165–166 °C. and from methanol, m.p. 213–214 °C. The product gave a positive Millon's reaction in the cold and in methanol gave a rose-red colour with a trace of ferric chloride. For analysis the *l*-bebeerine was dried at 117 °C. *in vacuo* for 3 hours.

Found: C, 72.0, 71.9, 72.5; H, 6.5, 6.5, 6.8; N, 5.0, 4.8; MeO, 10.4%.

Calculated for  $C_{36}H_{38}O_6N_2$ : C, 72.7; H, 6.4; N, 4.7; 2MeO, 10.4%.

(iii) *Separation of d-isoChondrodendrine*.—The residue from above was extracted with chloroform. The dark solution, after dilution with the same solvent, was then passed through a column of alumina which removed most of the pigments, giving a pale yellow chloroform eluate which was evaporated to dryness. The residue was dissolved in a mixture of methanol and chloroform, and evaporated to small bulk, when colourless microneedles separated, m.p. 300–302 °C. (decomp.) (5 g.). It gave a positive Millon's reaction in the cold.

Found: C, 71.7; H, 6.5; N, 4.8; MeO, 10.8%.

Calculated for  $C_{36}H_{38}O_6N_2$ : C, 72.7; H, 6.4; N, 4.7; 2MeO, 10.4%.

(iv) *l-Bebeerine Hydrochloride* crystallized when bebeerine was dissolved in hydrochloric acid, m.p. 270 °C. (decomp.).

(v) *l-Bebeerine Methiodide* was obtained by allowing the alkaloid to stand overnight with excess methyl iodide in methanol. On concentrating the solution the quaternary salt crystallized out, m.p. 250 °C. (decomp.).

(vi) *l-Bebeerine Dimethyl Ether*.—*l*-Bebeerine was allowed to stand overnight in the presence of excess ethereal diazomethane. After extracting with dilute sodium hydroxide, the ether was evaporated, but the residue could not be crystallized, nor could its salts.

Found: MeO, 18.1%.

Calculated for  $C_{38}H_{42}O_6N_2$ : 4MeO, 19.3%.

(vii) *O-Methyl Bebeerilene*.—The above dimethyl ether was converted to the methiodide and two successive Hofmann degradations(3) gave the nitrogen-free product, which on recrystallization from glacial acetic acid had m.p. 196 °C.

Found: C, 76.8; H, 5.8%.

Calculated for  $C_{36}H_{32}O_6$ : C, 77.1; H, 5.7%.

(viii) *d-isoChondrodendrine Sulphate*.—This salt crystallized readily from water in double pyramids which effloresced on standing and when dry had m.p. 285 °C. (decomp.).

(ix) *d-isoChondrodendrine Dimethyl Ether*.—The alkaloid, dissolved in a mixture of chloroform and methanol, was allowed to react overnight with excess ethereal diazomethane. The residue left after evaporation of the solvent was crystallized from benzene-light petroleum, m.p. 270–272 °C.

Found: C, 72.5; H, 6.8; N, 4.7; MeO, 19.4%.

Calculated for  $C_{38}H_{42}O_6N_2$ : C, 73.1; H, 6.8; N, 4.5; 4MeO, 19.3%.

(x) *d-isoChondrodendrine Dimethyl Ether Dimethiodide*.—The above methyl ether was allowed to stand for 1 hour with an excess of methyl iodide in methanol solution. The quaternary salt which separated was recrystallized from methanol, m.p. 310–312 °C.

Found: I, 27.3%.

Calculated for  $C_{40}H_{46}O_6N_2I_2$ : I, 28.0%.

(xi)  $\alpha$ -O-Methyl-isoChondrodendrine Methine.—This was prepared from the salt above by the method of King(2). The base had m.p. 205–206 °C. and dissolved in concentrated sulphuric acid to give a red solution which changed to indigo-blue on warming.

### III. ACKNOWLEDGMENTS

The authors are indebted to Miss J. Fildes and Miss B. Naylor, micro-analysts, University of Sydney, for all the analyses and to the Dunlop Rubber Company for the award of a scholarship to one of them (F.A.L.A.).

### IV. REFERENCES

- (1) DUTCHER, J. D.—*J. Amer. Chem. Soc.* **68** : 419 (1946).
- (2) KING, H.—*J. Chem. Soc.* **1940** : 737 (1940).
- (3) KING, H.—*J. Chem. Soc.* **1935** : 1386 (1935).
- (4) SPÄTH, E., LEITHE, W., and LADECK, F.—*Ber. dtsh. chem. Ges.* **61** : 1698 (1928).



## A CORRECTION

"Solar Radiation at 1200 Mc/s., 600 Mc/s., and 200 Mc/s.", by F. J. Lehany and D. E. Yabsley.  
*Aust. J. Sci. Res. A*, Vol. 2, No. 1, pp. 48-52, March 1949

Since publication of the above-mentioned paper, the aerial mounting has been improved so that the sun may be observed continuously. This has enabled the aerial focus to be checked more accurately using the sun itself as a distant source of signal instead of the previous arrangement using an oscillator approximately 100 feet from the aerial. This has shown that the original adjustment was in error and that, in consequence, the values of apparent solar temperature and flux densities of radiation recorded in the paper are too low. To correct for this error in focusing, these observations should be multiplied by the following factors.

Frequency (Mc/s.)	Multiplying Factor
1200	1.9
600	1.2
200	1.0

It will be necessary to apply the same correction factors to the 1200 and 600 Mc/s. data reported in the paper "Solar Radiation of Thermal Origin", by J. L. Pawsey and D. E. Yabsley, *Aust. J. Sci. Res. A*, Vol. 2, No. 2, pp. 198-213, June 1949.

With regard to the absolute error of the corrected measurements, the discussion on aerial directivity previously given in the paper still applies. The estimate of directivity is almost certainly an upper limit. Apart from this uncertainty, all other contributions to the absolute error in the observations are believed to amount to less than 20 per cent. Systematic errors, including that of aerial directivity, do not affect the relative values which are considered to be accurate to within 10 per cent.

F. J. Lehany  
D. E. Yabsley





# EXPERIMENTAL DESIGNS BALANCED FOR PAIRS OF RESIDUAL EFFECTS

By E. J. WILLIAMS\*

[*Manuscript received April 17, 1950*]

## *Summary*

When different treatments are applied in succession to the same unit of experimental material, it is necessary to consider the residual effects of preceding treatments on the present treatment. To facilitate the interpretation of the results of such experiments, balanced designs have been developed. In the designs discussed in this paper, each treatment is preceded by each pair of other treatments, so that the designs are balanced for the effects of pairs of residual effects.

It is shown that balanced designs are possible for  $n$  treatments, using  $n(n-1)$  replications. Solutions of the combinatorial problem involved can be derived from solutions of the Round Table problem, discussed in many books of mathematical puzzles.

The method of analysis of designs of this type is described, and exemplified by application to a set of experimental data.

## I. INTRODUCTION

In a recent paper(1), two types of experimental design, balanced with respect to the residual effect of treatments on those which followed them, were discussed. In such experiments, the treatments are applied successively on the same sites, in such a way that the residual effects of any treatment affect every other treatment equally. For the first type of experiment, only the residual effect of the immediately preceding treatment was considered. In a balanced design of this type, every treatment is preceded equally often by each other treatment. For the second type of experiment, the residual effects of two preceding treatments were considered, and balanced designs were found in which every treatment was preceded equally often, both immediately and by two positions, by each other treatment. Designs of the first type were found, for any number of treatments, but for the second type, only treatment numbers which were primes or powers of primes were considered. These latter solutions were derived from sets of mutually orthogonal Latin squares.

In this paper, designs which are balanced for pairs of residual effects are considered in more detail. Some such designs are based on sets of Latin squares, but others are not, and it appears that there is a large number of different possible designs even for a small number of treatments.

\* Section of Mathematical Statistics, C.S.I.R.O. ; working at Division of Forest Products, C.S.I.R.O., Melbourne.



The notation and terminology used in this paper follow closely those used in the previous paper. Residual effects of the immediately preceding, and the next preceding, treatments will be called first and second residual effects respectively.

## II. CONDITIONS FOR A BALANCED DESIGN

We take this opportunity to correct an error in the previous paper(1). A closer examination of the example given on page 155 (a design with 12 replications of 4 treatments), has revealed that, even though every treatment is preceded equally often both immediately and by two positions, by each other treatment, this design is not balanced for the estimation of pairs of residual effects. Direct effects are balanced with respect to each of the two residual effects, but the first and second residual effects are not balanced with respect to each other.

The reason for this is that, for a balanced design of this type, besides the conditions given previously, it is also necessary that each ordered pair of treatments occurs at the end of a row, equally often.

It can readily be seen that a design based on a set of orthogonal Latin squares, with the same initial columns, satisfies this condition, by the combinatorial theorem proved in the previous paper. Also, a design in which every ordered trio occurs equally frequently must have each ordered pair at the end of a row equally frequently. Designs other than these two types are not considered here.

## III. FIELD SOLUTIONS AND DIFFERENCE SOLUTIONS

If there are  $n$  treatments and each treatment is applied once on each site, the number of sites required for a balanced design is  $n(n-1)$ . Certain solutions may be arranged in sets of  $(n-1)$  Latin squares, in which the rows represent the sites, columns the order of application to the site, and the symbols the treatments.

When  $n$  is a prime or prime power, say  $p^k$ , it has been shown that a solution of the design problem may be constructed, each treatment being represented by one of the elements of the finite field  $GF(p^k)$ . These solutions, examples of which are given in (1), will be called *field solutions*.

Whether  $n$  is a prime power or not, each treatment may be represented by one of the classes of residues, modulo  $n$ . The differences between the successive treatments on any site are then considered. No difference may be zero, nor may the cumulative sum of any set of successive differences be zero. The condition that every treatment be preceded by every ordered pair of other treatments is equivalent to the condition that every treatment be preceded by every admissible ordered pair of differences (i.e. every ordered pair whose sum is not zero). There are  $(n-1)(n-2)$  such admissible pairs. For example, when  $n=5$ , the admissible pairs are

11	21	31	42
12	22	33	43
13	24	34	44

Any arrangement of the  $n$  different residues in a row gives  $(n-1)$  differences, which provide  $(n-2)$  pairs of adjacent differences. In a cyclic Latin square with such a set as its first row, each such pair of differences will precede each residue once. A set of  $n-1$  Latin squares, constructed in this way, which between them involve all the  $(n-1)(n-2)$  ordered pairs of differences, provides a solution of the problem. These solutions will be called *difference solutions*.

It will be noted that, for every ordered pair of treatments to occur at the end of a row, it is necessary that each difference occur once at the end of a row of differences.

The existence of difference solutions for general values of  $n$  was realized from a study of Dudeney's(2) Problem 273 (The Round Table) (see also Bergholt 3). In this problem, the  $n+1$  symbols are to be arranged in a circle, so that in  $\frac{1}{2}n(n-1)$  such arrangements each symbol is placed between each pair of other symbols once. One of Dudeney's methods for deriving solutions when  $n$  is greater than 4 is effectively to represent  $n$  of the symbols by the set of residue-classes, modulo  $n$ , and the remaining symbol by an element  $\infty$  which is unchanged by the addition to it of any other element. A set of initial arrangements  $\frac{1}{2}(n-1)$  or  $\frac{1}{2}n$  in number according as  $n$  is odd or even, is devised, and the remaining arrangements developed by adding 1, 2, . . . in turn to each element of an initial arrangement. It is shown that, when  $n$  is even, one of the cycles of arrangements thus produced consists of  $\frac{1}{2}n$  pairs of arrangements which are the reverse of each other. Only half this cycle is therefore required in the solution.

The solutions given in the previous paper(1), using finite fields, all provide solutions of the Round Table problem. For it can be readily seen, from the method of their construction, that the initial and final elements of a row take every pair of values once. Hence, if the ends of each row are joined up, an element  $\infty$  being placed between the initial and final element, the  $\infty$  element will occur once between each pair of other elements. It follows that every ordered trio occurs once in some cycle, thus fulfilling the conditions of the problem.

It will be seen that, when  $n$  is a prime, every field solution is a difference solution. However, difference solutions, which are not isomorphic with field solutions, exist, and will be illustrated.

When  $n$  is a power of a prime, the field solutions involve Latin squares which are non-cyclic, and are therefore not isomorphic with difference solutions.

#### IV. ENUMERATION OF DIFFERENCE SOLUTIONS

The enumeration of difference solutions consists of finding all possible arrangements of the  $(n-1)(n-2)$  admissible pairs of differences into  $n-1$  rows, each of  $n-1$  differences, such that sets of cyclic Latin squares can be derived from them. There are no difference solutions for  $n=4$ , and two non-isomorphic solutions for  $n=5$ . For  $n=6$ , six non-isomorphic solutions have been found, but others may exist.

TABLE 1  
DIFFERENCE SOLUTIONS BALANCED FOR PAIRS OF RESIDUAL EFFECTS

$n$	Solution Number	Initial Row	Sets of Differences
4	—	No solutions exist	
5	I*	0 1 4 2 3	1 3 3 1
		0 2 3 4 1	2 1 1 2
		0 3 2 1 4	3 4 4 3
		0 4 1 3 2	4 2 2 4
	II	0 1 2 4 3	1 1 2 4
		0 2 3 1 4	2 1 3 3
		0 3 2 4 1	3 4 2 2
		0 4 3 1 2	4 4 3 1
	I†	0 1 2 5 3 4	1 1 3 4 1
		0 2 3 5 4 1	2 1 2 5 3
		0 3 4 2 1 5	3 1 4 5 4
		0 4 2 5 1 3	4 4 3 2 2
		0 5 1 4 3 2	5 2 3 5 5
	II	0 1 2 5 4 3	1 1 3 5 5
		0 2 3 5 4 1	2 1 2 5 3
		0 3 4 2 1 5	3 1 4 5 4
		0 4 2 5 1 3	4 4 3 2 2
		0 5 1 4 2 3	5 2 3 4 1
	III	0 1 2 5 4 3	1 1 3 5 5
		0 2 1 4 5 3	2 5 3 1 4
		0 3 1 2 4 5	3 4 1 2 1
		0 4 2 5 1 3	4 4 3 2 2
		0 5 3 2 4 1	5 4 5 2 3
	IV	0 1 3 4 5 2	1 2 1 1 3
		0 2 1 4 5 3	2 5 3 1 4
		0 3 2 1 5 4	3 5 5 4 5
		0 4 2 5 1 3	4 4 3 2 2
		0 5 1 4 2 3	5 2 3 4 1
	V	0 1 2 5 3 4	1 1 3 4 1
		0 2 1 4 5 3	2 5 3 1 4
		0 3 5 1 2 4	3 2 2 1 2
		0 4 3 1 5 2	4 5 4 4 3
		0 5 1 4 3 2	5 2 3 5 5
	VI	0 1 2 5 4 3	1 1 3 5 5
		0 2 1 4 5 3	2 5 3 1 4
		0 3 5 1 2 4	3 2 2 1 2
		0 4 3 1 5 2	4 5 4 4 3
		0 5 1 4 2 3	5 2 3 4 1

\* Field solution.

† Isomorphic with solution given by Bergholt(3) and Dudeney(2).

The known solutions, for  $n=5$  and 6, are given in Table 1. The initial rows of the  $n-1$  Latin squares are tabulated, and also the rows of differences from which they are derived.

### V. OTHER SOLUTIONS

With 6 treatments, a balanced design requires 30 replications, or 180 applications of treatment in all. For this reason it is unlikely that more than 6 treatments would be used in practice. It is not proposed, therefore, to discuss the various possible designs for values of  $n$  greater than 6. It may be noted, however, that Dudeney's solutions(2) of the Round Table problem provide solutions of the design problem for values of  $n$  up to and including 11. His solutions for  $n=7, 11$ , are equivalent to the solution given previously(1) for all prime  $n$ . His solution for  $n=9$  is a difference solution. His solutions for  $n=8, 10$ , are of a new type, neither field solutions nor difference solutions. In fact, they are not even sets of Latin squares. The existence of these solutions indicates that there may be several methods of constructing such designs other than those so far discussed.

Of these other types of designs it suffices to mention only an example for 5 treatments. This has been derived from one of Bergholt's solutions(3). Here the treatments are designated by members of the set of residue-classes (modulo 4) together with an element  $\infty$  which is unchanged by the addition to it of any other element. The 20 replications are made up of 5 cycles of 4, of which the initial rows are

$\infty$	0	2	1	3
0	$\infty$	3	2	1
0	3	$\infty$	1	2
0	1	2	$\infty$	3
0	2	3	1	$\infty$ .

(Thus the first cycle is

$\infty$	0	2	1	3
$\infty$	1	3	2	0
$\infty$	2	0	3	1
$\infty$	3	1	0	2.)

### VI. ANALYSIS OF RESULTS

The analysis of experiments, taking into account a pair of residual effects, is more complicated than that for a single residual effect. Also, the method of analysis will differ, according as the interactions of pairs of residual effects are or are not considered. Only the analysis ignoring such interactions will be considered here. This is in line with the treatment in the previous paper, where the interaction of residual and direct effects was ignored. This is necessary in the case where  $n$  replications are used for  $n$  treatments, for then the "error" variation is in fact attributable to the interaction of direct and residual effects.

The analysis taking into account pairs of residual effects will be considered for  $n(n-1)$  replications of  $n$  treatments. The following notation will be used.



$$\begin{aligned}
 & \left. \begin{array}{l} \text{Totals} \end{array} \right\} \begin{cases} T_i = \text{total of results for treatment } i, \\ R_i = \text{total of results for treatments preceded immediately by} \\ \quad \text{treatment } i, \\ Q_i = \text{total of results for treatments preceded by two positions by} \\ \quad \text{treatment } i, \\ F_i = \text{total of results for all sites on which treatment } i \text{ is the final} \\ \quad \text{one,} \\ E_i = \text{total of results for all sites on which treatment } i \text{ is the last} \\ \quad \text{but one,} \\ P_j = \text{total of results for position } j \text{ in order of application to the} \\ \quad \text{site,} \\ G = \text{grand total of all results,} \\ R'_i = R_i + \frac{P_1}{n}, \\ Q'_i = Q_i + \frac{P_1 + P_2}{n}. \end{cases} \\
 \\
 & \left. \begin{array}{l} \text{Estimates} \\ \text{(measured} \\ \text{as devia-} \\ \text{tions from} \\ \text{zero)} \end{array} \right\} \begin{cases} t_i = \text{direct effect of treatment } i, \\ r_i = \text{first residual effect of treatment } i, \\ q_i = \text{second residual effect of treatment } i, \\ f_i = \text{average effect for all sites on which treatment } i \text{ is the} \\ \quad \text{final one,} \\ e_i = \text{average effect for all sites on which treatment } i \text{ is the} \\ \quad \text{last but one,} \\ p_j = \text{effect of position } j \text{ in order of application.} \end{cases} \\
 & \quad i, j = 1, 2 \dots n.
 \end{aligned}$$

(a) *Derivation of the Estimates, and Tests of Significance*

Since the effect of position in order of application is orthogonal with all the other effects, it need not be considered when the estimates of the effects of other factors are being determined. For the estimates of position effect we have simply

$$n(n-1)p_j = P_j - \frac{G}{n}.$$

For the remaining effects, the normal equations are most conveniently set out in matrix form. They are

$$\begin{bmatrix} n(n-1) & -(n-1) & -(n-2) & 0 & 0 \\ -(n-1) & (n-1)^2 & -(n-2) & -(n-1) & 1 \\ -(n-2) & -(n-2) & (n-1)(n-2) & -(n-2) & -(n-2) \\ 0 & -(n-1) & -(n-2) & n(n-1) & -n \\ 0 & 1 & -(n-2) & -n & n(n-1) \end{bmatrix} \begin{bmatrix} t_i \\ r_i \\ q_i \\ f_i \\ e_i \end{bmatrix} = \begin{bmatrix} T_i - \frac{G}{n} \\ R'_i - \frac{G}{n} \\ Q'_i - \frac{G}{n} \\ F_i - \frac{G}{n} \\ E_i - \frac{G}{n} \end{bmatrix}$$

The solutions of these equations are determined by the use of the inverse matrix, which is found to be

$$\frac{1}{n^3 - n^2 - 5n - 2} \begin{bmatrix} \frac{(n+1)(n^2-2n-1)}{n(n-1)} & \frac{n+1}{n-1} & \frac{n}{n-2} & \frac{2(n^2-n-1)}{n(n-1)(n-2)} & \frac{1}{n-2} \\ \frac{n+1}{n-1} & \frac{n^2-2}{n-1} & \frac{n+2}{n-2} & \frac{n^3-n^2-n+2}{n(n-1)(n-2)} & \frac{n+2}{n(n-2)} \\ \frac{n}{n-2} & \frac{n+2}{n-2} & \frac{n^2-1}{n-2} & \frac{n^2+n+1}{n(n-2)} & \frac{n^3-1}{n(n-2)} \\ \frac{2(n^2-n-1)}{n(n-1)(n-2)} & \frac{n^3-n^2-n+2}{n(n-1)(n-2)} & \frac{n^2+n+1}{n(n-2)} & \frac{n^5-3n^4+6n^3-2n-1}{n^2(n-1)(n-2)} & \frac{n^3-4n-1}{n^2(n-2)} \\ \frac{1}{n-2} & \frac{n+2}{n(n-2)} & \frac{n^2-1}{n(n-2)} & \frac{n^3-4n-1}{n^2(n-2)} & \frac{(n-1)(n^3-n^2-4n-1)}{n^2(n-2)} \end{bmatrix}$$

In practice, the estimates required will be those for each treatment effect (direct or residual) with and without adjustment for the other treatment effects. Thus, the estimates of each of the treatment effects (direct and residual), adjusted for both the other effects, will be needed to provide a test for their significance. Also, estimates of direct effects, adjusted for the first residual effect alone, and unadjusted, will be required, according as one or none of the residual effects is found to be significant. As it is unlikely that second residual effects would exist when first residual effects were absent, the adjustment of direct effects for second residual effects alone need not be considered.

Corresponding to the different estimates there will be different sums of squares in the analysis of variance. There are six possible ways of apportioning the sum of squares for treatments among direct effects and first and second residual effects, but not all are of practical interest. The sums of squares will need to be apportioned in different ways, in order to test the significance of each treatment effect. The calculations can be so arranged that each partition of the total sum of squares for treatments involves different treatment effects. As the sum of squares for each such effect may be calculated directly, they provide an accurate check on the calculations.

The estimates of the different treatment effects may be obtained by modifying the original set of normal equations. Where an effect is to be ignored, it is simply omitted from the equations. To distinguish the different estimates, the subscript  $i$  representing the  $i$ th treatment will be dropped, and the following notation used:

- $t$  = direct effect, unadjusted,
- $r$  = first residual effect, unadjusted,
- $q$  = second residual effect, unadjusted,
- $t_r$  = direct effect, adjusted for first residual effect,
- $r_t$  = first residual effect, adjusted for direct effect,
- $r_q$  = first residual effect, adjusted for second residual effect,
- $q_t$  = second residual effect, adjusted for direct effect,
- $t_{rq}$  = direct effect, adjusted for first and second residual effects,
- $r_{tq}$  = first residual effect, adjusted for direct and second residual effects,
- $q_{tr}$  = second residual effect, adjusted for direct and first residual effects.

Then the various estimates are found to be as follows. The unadjusted estimates are

$$\begin{aligned} n^2(n-1)t &= nT_i - G \\ &= T, \end{aligned}$$

$$\begin{aligned} n(n-1)(n^2-n-1)r &= n^2R'_i + nF_i - (n+1)G \\ &= R, \end{aligned}$$

$$\begin{aligned} (n+1)n(n-2)^2q &= n^2Q'_i + nF_i + nE_i - (n+2)G \\ &= Q. \end{aligned}$$

Estimates adjusted for one other effect are

$$\begin{aligned} (n+1)n(n-1)(n-2)t_r &= (n^2-n-1)T_i + nR'_i + F_i - nG \\ &= T_r, \end{aligned}$$

$$\begin{aligned} (n+1)n(n-1)(n-2)r_i &= nT_i + n^2R'_i + nF_i - (n+2)G \\ &= R_i, \end{aligned}$$

$$\begin{aligned} n(n-2)(n^2-2n-1)r_q &= n(n-2)R'_i + nQ'_i + (n-1)F_i + E_i - nG \\ &= R_q, \end{aligned}$$

$$\begin{aligned} n(n-2)^2(n^2-2)q_i &= n(n-2)T_i + n^2(n-1)Q'_i + n(n-1)F_i \\ &\quad + n(n-1)E_i - (n^2+2n-4)G \\ &= Q_i. \end{aligned}$$

Finally, using the inverse matrix given above, the completely adjusted estimates are found to be

$$\begin{aligned} n(n-1)(n-2)(n^3-n^2-5n-2)t_{rq} &= (n+1)(n-2)(n^2-2n-1)T_i + (n+1)n(n-2)R'_i \\ &\quad + n^2(n-1)Q'_i + 2(n^2-n-1)F_i + n(n-1)E_i \\ &\quad - n^2(n-1)G \\ &= T_{rq}. \end{aligned}$$

$$\begin{aligned} n(n-1)(n-2)(n^3-n^2-5n-2)r_{iq} &= (n+1)n(n-2)T_i + n(n-2)(n^2-2)R'_i \\ &\quad + n(n-1)(n+2)Q'_i + (n^3-n^2-n+2)F_i \\ &\quad + (n-1)(n+2)E_i - n(n-1)(n+2)G \\ &= R_{iq}. \end{aligned}$$

$$\begin{aligned} n(n-2)(n^3-n^2-5n-2)q_{tr} &= n^2T_i + n(n+2)R'_i + (n+1)n(n-1)Q'_i \\ &\quad + (n^2+n+1)F_i + (n^2-1)E_i - (n^2+4n+2)G \\ &= Q_{tr}. \end{aligned}$$

The rather lengthy expressions for the completely adjusted estimates may be calculated quite simply by a systematic method, which yields as by-products the other estimates required. The estimates of increasing complexity are built up successively from the less complex ones as shown in Table 2.

In the analysis of variance, the required sums of squares for the different effects may be determined by calculating the sums of squares of the different quantities given in the table, and dividing by an appropriate divisor; these divisors are tabulated alongside the expressions to be squared, in Table 2.

TABLE 2

QUANTITIES TO BE CALCULATED IN DETERMINING ESTIMATES OF TREATMENT EFFECTS, AND SUMS OF SQUARES FOR TREATMENTS IN THE ANALYSIS OF VARIANCE

Quantity	Calculated from	Divisor for Sum of Squares
$Q$	—	$(n+1)n^2(n-2)^2$
$R$	—	$n^3(n-1)(n^2-n-1)$
$T$	—	$n^3(n-1)$
$Q_t$	$(n-2)T + (n-1)Q$	$n^3(n-1)(n-2)^2(n^2-2)$
$R_q$	$1/n\{(n-2)R + Q\}$	$n^2(n-2)^2(n^2-2n-1)$
$R_t$	$T + R$	$(n+1)n^3(n-1)(n-2)$
$T_r$	$1/n\{(n+1)(n-2)T + R_r\}$	$(n+1)n(n-1)(n-2)(n^2-n-1)$
$Q_{tr}$	$1/n\{(n+2)R_t + (n+1)Q_t\}$	$(n+1)n^2(n-1)(n-2)(n^3-n^2-5n-2)$
$R_{tq}$	$1/n\{(n^2-2)R_t + (n+2)Q_t\}$	$n^2(n-1)(n-2)^2(n^2-2)(n^3-n^2-5n-2)$
$T_{rq}$	$(n+1)(n-2)T_r + Q_t - (n-2)(n+2)T$	$(n+1)n(n-1)(n-2)^2(n^2-2n-1)(n^3-n^2-5n-2)$

From the quantities given in Table 2, the sum of squares for treatments may be determined in four different ways. These correspond to the following partitions of treatment effects:

1. Direct effects, adjusted for first and second residual effects ( $t_{rq}$ )  
 First residual effects, adjusted for second residual effects ( $r_q$ )  
 Second residual effects, unadjusted ( $q$ )
2. First residual effects, adjusted for direct and second residual effects ( $r_{tq}$ )  
 Second residual effects, adjusted for direct effects ( $q_t$ )  
 Direct effects, unadjusted ( $t$ )
3. Second residual effects, adjusted for direct and first residual effects ( $q_{tr}$ )  
 Direct effects, adjusted for first residual effects ( $t_r$ )  
 First residual effects, unadjusted ( $r$ )
4. Second residual effects, adjusted for direct and first residual effects ( $q_{tr}$ )  
 First residual effects, adjusted for direct effects ( $r_t$ )  
 Direct effects, unadjusted ( $t$ )

The first three analyses enable the significance of each effect to be tested, as well as giving a complete check on the calculations. The third and fourth analyses are needed to test the significance of direct and first residual effects, in case second residual effects are to be ignored.

Apart from the various ways of breaking up the treatment sum of squares, the analysis of variance takes the usual form, as can be seen from the example given below.



*(b) Accuracy of the Estimates*

If  $E$  is the error mean square from the analysis of variance, the estimated variances of the different estimates are given as follows:

Estimate	Estimated Variance
$t$	$\frac{E}{n(n-1)}$
$r$	$\frac{nE}{(n-1)(n^2-n-1)}$
$q$	$\frac{nE}{(n+1)(n-2)^2}$
$t_r$	$\frac{(n^2-n-1)E}{(n+1)n(n-1)(n-2)}$
$r_t$	$\frac{nE}{(n+1)(n-1)(n-2)}$
$r_q$	$\frac{E}{n^2-2n-1}$
$q_t$	$\frac{n(n-1)E}{(n-2)^2(n^2-2)}$
$t_{rq}$	$\frac{(n+1)(n^2-2n-1)E}{n(n-1)(n^3-n^2-5n-2)}$
$r_{tq}$	$\frac{(n^2-2)E}{(n-1)(n^3-n^2-5n-2)}$
$q_{tr}$	$\frac{(n+1)(n-1)E}{(n-2)(n^3-n^2-5n-2)}$

The main use of these estimates of variance is to enable the differences between pairs of values to be tested for significance.

*(c) Application to a Numerical Example*

Experimental data exemplifying the designs given above are presented by Cochran, Autrey, and Cannon(4) in their work on designs balanced for the estimation of residual effects. In their experiment, 18 cows were each fed three different rations, the effect studied being that of ration on milk yield. The experiment was designed to be balanced for one residual effect, but with three treatments, two  $3 \times 3$  Latin squares are required, and these are balanced for both residual effects. The basic design requires six cows, so that the design is replicated three times in the experiment.

The data from this experiment are here analysed to illustrate the method of analysis for pairs of residual effects. One of the cows in the experiment was lost, and the missing values were fitted in the original analysis. These values are used in the present analysis in order to make the results comparable. The notation used is that of the present paper, and not that of (4).

TABLE 3  
TOTALS OF FAT CORRECTED MILK UNITS FROM A CATTLE FEEDING EXPERIMENT

$i$	$T_i$	$R'_i$	$Q'_i$	$F_i$	$E_i$	$j$	$P_j$
1	22,875	25,887	26,216.3	26,226	25,598	1	29,811
2	25,926	25,344	25,600.3	26,294	24,886	2	25,837
3	28,092	25,662	25,076.3	24,373	26,409	3	21,245
Total ..	76,893	76,893	76,893	76,893	76,893	Total ..	76,893

The original milk yields are given in (4), and are not reproduced here. In Table 3, the various totals required in the analysis are set out. In Table 4, the estimates of the various effects, adjusted and unadjusted, are set out. The unadjusted estimates of the treatment effects are found directly from Table 3; thus, for treatment 1,

$$\begin{aligned}
 Q &= 9Q'_1 + 3F_1 + 3E_1 - 5G \\
 &= 9 \times 26,216.3 + 3 \times 26,226 + 3 \times 25,598 - 5 \times 76,893 \\
 &= 6,954
 \end{aligned}$$

and similarly for the other estimates, using the formulae given above. The adjusted estimates are derived from the unadjusted estimates, using the relations given in the third column of Table 4. As an example, for treatment 1,

$$\begin{aligned}
 T_{rq} &= 4 \times -12,417 + 5,640 - 5 \times -8,268 \\
 &= -2,688.
 \end{aligned}$$

TABLE 4  
CALCULATIONS OF ESTIMATES OF TREATMENT EFFECTS, AND SUMS OF SQUARES FOR TREATMENTS IN THE ANALYSIS OF VARIANCE (CF. TABLE 2)

Quantity	In Terms of Estimates* $3 \times$	Calculated from	Treatments			Divisor for Sum of Squares* $3 \times$	Sum of Squares
			1	2	3		
$Q$	$12q$	—	6,954	-522	-6,432	108	277,782
$R$	$30r$	—	4,089	-594	-3,495	270	36,158
$T$	$18t$	—	-8,268	885	7,383	54	763,282
$Q_t$	$21q_t$	$T + 2Q$	5,640	-159	-5,481	378	54,565
$R_q$	$6r_q$	$1/3(R + Q)$	3,681	-372	-3,309	18	456,252
$R_t$	$12r_t$	$T + R$	-4,179	291	3,888	216	50,409
$T_r$	$24t_r$	$1/3(4T + R_t)$	-12,417	1,277	11,140	120	777,534
$Q_{tr}$	$3q_{tr}$	$1/3(5R_t + 4Q_t)$	555	273	-828	72	4,945
$R_{tq}$	$6r_{tq}$	$1/3(7R_t + 5Q_t)$	-351	414	-63	126	790
$T_{rq}$	$6t_{rq}$	$4T_r + Q_t - 5T$	-2,688	524	2,164	48	84,603

\* The factor 3 is introduced because there are three replications of the basic design.

TABLE 5  
ANALYSES OF VARIANCE

(a) *Main Analysis*

	Degrees of Freedom	Sum of Squares	Mean Square
Groups .. .. .	5	1,311,769	262,354
Cows within groups .. .. .	11	654,638	59,513
Periods .. .. .	2	2,041,769	
Group-period interaction .. .. .	10	193,341	19,334
Rations (direct and both residual effects) .. .. .	6	818,637	136,440
Error .. .. .	16	104,113	6,507
Total .. .. .	50	5,124,267	

(b) *Sub-Analyses of Treatment Effects*

	Corresponding to	Degrees of Freedom	Sum of Squares	Mean Square
(i)	$q$ and $r_q$	4	734,034	42,301**
	$t_{rq}$	2	84,603	
(ii)		6	818,637	395
	$t$ and $q_t$	4	817,847	
(iii)		2	790	388,767** 2,473
	$r$	2	36,158	
	$t_r$	2	777,534	
(iv)	$y_{tr}$	2	4,945	25,205* 2,473
		6	818,637	
	$t$	2	763,282	
	$r_t$	2	50,409	
	$q_{tr}$	2	4,945	
		6	818,636	

\* Significant at the 5% level.

\*\* Significant at the 1% level.

If the actual estimates of treatment effects are required, they may be obtained by dividing the quantities in columns 4, 5, and 6 of Table 4 by the coefficients in column 2. A factor 3 has to be included in this divisor since there are three replications of the basic design.

The final two columns of Table 4 give the quantities required for calculating the analysis of variance. The sum of squares of corresponding figures in columns 4, 5, and 6, divided by the figure in column 7, gives the appropriate sum of squares for the analysis of variance, shown in column 8. Thus, the sum of squares for direct effects, adjusted for both residual effects, is

$$\frac{1}{3 \times 48} \left\{ (-2,688)^2 + 524^2 + 2,164^2 \right\} = 84,603.$$

The analysis of variance is given in Table 5, together with the sub-analyses required for testing the significance of direct and residual effects. The direct effect of treatments is still highly significant, after adjustment for both residual effects (i). The second residual effect is not significant (iii), (iv), and it may therefore be ignored. On this basis, the first residual effect attains significance at the 5 per cent. level (iv), and the direct effect at the 1 per cent. level (iii).

While the second residual effect is not significant, as could well be expected in this example, the analysis set out shows how significance is to be tested, and how adjustment may be made for the second residual effect if it appears necessary to do so.

## VII. REFERENCES

- (1) WILLIAMS, E. J.—Experimental designs balanced for the estimation of residual effects of treatments. *Aust. J. Sci. Res. A* **2**: 149 (1949).
- (2) DUDENEY, H. E.—“Amusements in Mathematics.” (Thomas Nelson and Sons Ltd.: London, 1943.)
- (3) ROUSE BALL, W. W.—“Mathematical Recreation.” Vol. 3, pp. 259, 290, 320, 430, citing Bergholt’s solutions (1906). (Macmillan & Co.: London, 1939.)
- (4) COCHRAN, W. G., AUTREY, K. M., and CANNON, C. Y.—A double change-over design for dairy cattle feeding experiments. *J. Dairy Sci.* **24**: 937 (1941).



# ESTIMATION OF THE ERRORS OF THE LEAST-SQUARES POLYNOMIAL COEFFICIENTS

By P. G. GUEST\*

[*Manuscript received March 24, 1950*]

## *Summary*

The estimation of the errors in the values obtained for the power-series and differential coefficients of a least-squares curve fitted to a number of equally-spaced observations is discussed. Curves and tables of the various weight functions are obtained.

## I. INTRODUCTION

In investigations where the aim in fitting a least-squares curve is merely to smooth the data, the actual form in which the curve is expressed is not important. In other cases, however, the coefficients of the power-series expansion may be of considerable significance. Thus the result required might be the least-squares curve expressed in power-series form, or it might be the values of one or more of the differential coefficients at certain values of the variable.

In the present paper the estimation of the errors of the differential and polynomial coefficients is discussed. The treatment is similar to that of a previous paper(1), in which the estimated errors of the fitted values were considered. Graphs and tables of the various weight functions are obtained from which the appropriate weight function may be read off directly. The tables may be used as a rapid alternative to Birge's method(2) for estimating the errors of these coefficients.

## II. CHANGE OF ORIGIN AND SCALE

In most methods of fitting polynomials to a series of  $n$  equally-spaced observations, the calculations are carried out in terms of the variable  $\varepsilon$  whose values at the points of observation are the integers or half-integers  $-\frac{1}{2}(n-1)$ ,  $-\frac{1}{2}(n-3)$ , ...,  $+\frac{1}{2}(n-1)$ .

Since in general the "physical" variable, which we shall denote by  $x$ , has neither the same origin nor the same scale as the variable  $\varepsilon$ , it is necessary to derive formulae for transferring from  $\varepsilon$  to  $x$ . Suppose that

$$x = h(\varepsilon - m), \quad \dots\dots\dots (1)$$

that is, the variable  $x$  has its origin at  $\varepsilon = m$  and its scale is  $h$  times as great as that of  $\varepsilon$ . (Note that both Birge(2) and Kerawala(3) use the transformation  $x = h\varepsilon + m$ . Thus  $m$  in our notation corresponds to  $-m/h$  in theirs.)

\* Physics Department, University of Sydney.

The least-squares polynomial of degree  $p$  may be written in the following forms :

$$u_p(\varepsilon) = b_0 + b_1 V_1(\varepsilon) + b_2 V_2(\varepsilon) + \dots + b_p V_p(\varepsilon), \dots \quad (2)$$

where  $V_j(\varepsilon)$  is the orthogonal polynomial of degree  $j$  in  $\varepsilon$ ;

$$u_p(\varepsilon) = b_{p0} + b_{p1}\varepsilon + b_{p2}\varepsilon^2 + \dots + b_{pp}\varepsilon^p; \dots \quad (3)$$

and

$$u_p(x) = c_{p0} + c_{p1}x + c_{p2}x^2 + \dots + c_{pp}x^p. \dots \quad (4)$$

$c_{pj}$  can be expressed either in terms of the coefficients  $b_q$  or the coefficients  $b_{pq}$ . We have in the first case

$$c_{pj} = h^{-j} \sum_{q=j}^p H_{jq} b_q, \dots \quad (5)$$

where

$$H_{jq} = \sum_{r=j}^q m^{r-j} \binom{r}{j} S_{rq}, \dots \quad (6)$$

and  $S_{rq}$  is the coefficient of  $\varepsilon^r$  in the power-series expansion of  $V_q(\varepsilon)$ . In the second case

$$c_{pj} = h^{-j} \sum_{q=j}^p \left(\frac{q}{j}\right) m^{q-j} b_{pq}, \dots \quad (7)$$

If the non-orthogonal method of curve-fitting due to Kerawala(3) is used, the coefficients  $b_{pq}$  are obtained directly, and the  $c_{pj}$  may be obtained by means of a Horner shift or from equation (7). If orthogonal methods of curve-fitting are used, the  $b_{pq}$  must first be obtained from the coefficients  $b_j$  by means of the relation

$$b_{pq} = \sum_{r=q}^p S_{qr} b_r.$$

Thus with orthogonal methods the same amount of work is involved in performing a single shift of origin, whether equation (5) or equation (7) is used. The use of equation (7) is preferable when a number of shifts are to be performed, for example, when the differential coefficients are required for a number of values of the variable. However, the terms  $H_{jq}$  are used in Birge's scheme for the determination of probable errors.

### III. THE ERRORS OF THE COEFFICIENTS

The coefficients  $b_q$ , being orthogonal coefficients, are uncorrelated and of weight  $\sum_{\varepsilon} V_q^2$ . Writing  $N_q$  for  $\sum_{\varepsilon} V_q^2$ , it follows from (5) that the estimated error of the coefficient  $c_{pj}$  is

$$r[c_{pj}(m)] = h^{-j} \left( \sum_{q=j}^p H_{jq}^2 / N_q \right)^{\frac{1}{2}} r_p, \dots \quad (8)$$

where  $r_p$  is the estimated error of an observation.

The errors can be expressed either as standard errors or as probable errors, the formulae relating the estimated error of  $c_{pj}$  to the estimated error of an observation being of course identical in the two cases. For convenience we shall use the general term "estimated error" to denote either the estimated standard error or the estimated probable error.

The estimated errors of the coefficients  $b_{pj}$  and  $b_j$  may be obtained by substituting the values  $m=0$  and  $h=1$  in equation (8). Thus

$$r[b_{pj}] = r[c_{pj}(0)] \quad (h=1)$$

$$\text{and } r[b_j] = S_{jj}r[c_{jj}(0)] \quad (h=1).$$

The quantity  $S_{jj}$ —in the notation of Fisher(4),  $\lambda_j$ —is just the factor by which the Tchebycheff polynomial  $T_j(\varepsilon)$  is multiplied to give the integral values  $V_j(\varepsilon)$  or  $\xi'(\varepsilon)$ .

Introducing the variable  $k=2m/n$ , equation (8) can be put in the form

$$r[c_{pj}(k)] = \left\{ \sum_{q=j}^p \tau_{qj}(k, n) \right\}^{\frac{1}{2}} r_p / h^j n^{j+\frac{1}{2}}, \quad \dots\dots\dots (9)$$

$$= \rho_{pj}(k, n) r_p / h^j n^{j+\frac{1}{2}}, \text{ say. } \dots\dots\dots (10)$$

The quantities  $\tau, \rho$  can be written as explicit functions of  $k$  by using the standard expressions for the orthogonal polynomials. Thus we obtain the following functions :

$$\tau_{11}(k, n) = 12D_1$$

$$\tau_{21}(k, n) = 180k^2D_2$$

$$\tau_{31}(k, n) = 7\{15k^2 - (3 - 7v^2)\}^2D_3$$

$$\tau_{41}(k, n) = 225k^2\{7k^2 - (3 - 13v^2)\}^2D_4$$

$$\tau_{51}(k, n) = \frac{11}{16}\{315k^4 - 210(1 - 7v^2)k^2 + (15 - 230v^2 + 407v^4)\}^2D_5$$

$$\tau_{22}(k, n) = 180D_2$$

$$\tau_{32}(k, n) = 6300k^2D_3$$

$$\tau_{42}(k, n) = 225\{21k^2 - (3 - 13v^2)\}^2D_4$$

$$\tau_{52}(k, n) = 121275k^2\{3k^2 - (1 - 7v^2)\}^2D_5$$

$$\tau_{33}(k, n) = 2800D_3$$

$$\tau_{43}(k, n) = 176400k^2D_4$$

$$\tau_{53}(k, n) = 53900\{9k^2 - (1 - 7v^2)\}^2D_5$$

$$\tau_{44}(k, n) = 44100D_4$$

$$\tau_{54}(k, n) = 4365900k^2D_5$$

$$\tau_{55}(k, n) = 698544D_5$$

where

$$D_j = 1/(1 - v^2)(1 - 4v^2) \dots (1 - j^2v^2)$$

and

$$v^2 = 1/n^2.$$

The corresponding expressions  $\tau_{p0}$  have been given in a previous paper(1).

The range of values of the variable  $k$  may be divided into two parts: the region of interpolation ( $|k| < 1$ ) and the region of extrapolation ( $|k| > 1$ ).

In the region of interpolation the functions  $\rho_{pj}(k, n)$  are given in Table 1 for values of  $p$  up to 5, and for selected values of  $n$ . Intermediate values may be obtained by interpolating between the tabulated values. The error arising from interpolation between the tabulated values is usually considerably less

TABLE 1

$n \backslash k$		$\rho_{11}(k, n)$							
$n$		$\infty$	25	14	11	8	7		
$k$	all	3.46	3.47	3.47	3.48	3.49	3.50		
$n \backslash k$		$\rho_{21}(k, n)$							
$n$		$\infty$	25	14	11	8	7		
$k$									
0.00		3.46	3.47	3.47	3.48	3.49	3.50		
0.05		3.53	3.53	3.54	3.55	3.56	3.57		
0.10		3.71	3.72	3.73	3.74	3.76	3.78		
0.15		4.01	4.01	4.03	4.04	4.07	4.09		
0.20		4.38	4.39	4.41	4.43	4.47	4.50		
0.25		4.82	4.83	4.86	4.88	4.94	4.98		
0.30		5.31	5.32	5.36	5.38	5.45	5.50		
0.35		5.84	5.85	5.89	5.92	6.01	6.06		
0.40		6.39	6.41	6.45	6.49	6.59	6.65		
0.45		6.96	6.98	7.03	7.08	7.19	7.26		
0.50		7.55	7.58	7.63	7.68	7.81	7.89		
0.55		8.15	8.18	8.24	8.30	8.44	8.53		
0.60		8.76	8.79	8.86	8.93	9.08	9.18		
0.65		9.38	9.42	9.49	9.56	9.73	9.84		
0.70		10.01	10.05	10.13	10.20	10.38	10.50		
0.75		10.64	10.68	10.77	10.85	11.04	11.17		
0.80		11.3	11.3	11.4	11.5	11.7	11.8		
0.85		11.9	12.0	12.1	12.2	12.4	12.5		
0.90		12.6	12.6	12.7	12.8	13.0	13.2		
0.95		13.2	13.3	13.4	13.5	13.7	13.9		
1.00		13.9	13.9	14.0	14.1	14.4	14.6		
$n \backslash k$		$\rho_{31}(k, n)$							
$n$		$\infty$	40	25	18	14	11	8	7
$k$									
0.00		8.66	8.68	8.72	8.77	8.84	8.96	9.27	9.49
0.05		8.60	8.62	8.65	8.70	8.78	8.89	9.19	9.41
0.10		8.41	8.43	8.46	8.51	8.58	8.69	8.97	9.17
0.15		8.10	8.12	8.15	8.19	8.26	8.36	8.61	8.79
0.20		7.71	7.73	7.75	7.79	7.84	7.93	8.14	8.30
0.25		7.28	7.29	7.31	7.34	7.38	7.45	7.62	7.74
0.30		6.87	6.88	6.90	6.92	6.95	7.00	7.12	7.20
0.35		6.60	6.60	6.61	6.63	6.65	6.69	6.77	6.83
0.40		6.58	6.59	6.60	6.61	6.63	6.67	6.75	6.80
0.45		6.96	6.97	6.98	7.01	7.04	7.09	7.20	7.29
0.50		7.81	7.82	7.84	7.88	7.93	8.01	8.21	8.36
0.55		9.11	9.13	9.17	9.23	9.30	9.43	9.75	9.98
0.60		10.8	10.9	10.9	11.0	11.1	11.3	11.7	12.1
0.65		12.9	12.9	13.0	13.1	13.3	13.5	14.1	14.6
0.70		15.3	15.3	15.4	15.5	15.7	16.0	16.8	17.4
0.75		17.9	18.0	18.1	18.2	18.5	18.9	19.8	20.5
0.80		20.8	20.9	21.0	21.2	21.5	21.9	23.1	24.0
0.85		23.9	24.0	24.2	24.4	24.7	25.3	26.7	27.6
0.90		27.3	27.4	27.6	27.8	28.2	28.9	30.4	31.6
0.95		30.8	31.0	31.2	31.5	31.9	32.7	34.5	35.8
1.00		34.6	34.8	35.0	35.4	35.9	36.7	38.7	40.2



TABLE 1 (Continued)

$n$		$\rho_{41}(k, n)$							
$k$	$\infty$	40	25	18	14	11	9	8	7
0.00	8.7	8.7	8.7	8.8	8.8	9.0	9.1	9.3	9.5
0.05	8.9	8.9	8.9	9.0	9.1	9.2	9.4	9.6	9.8
0.10	9.5	9.5	9.6	9.6	9.8	9.9	10.2	10.4	10.8
0.15	10.3	10.4	10.4	10.5	10.7	10.9	11.2	11.6	12.1
0.20	11.2	11.3	11.4	11.5	11.7	11.9	12.4	12.7	13.4
0.25	12.1	12.1	12.2	12.3	12.5	12.9	13.3	13.8	14.5
0.30	12.7	12.7	12.8	13.0	13.2	13.5	14.1	14.5	15.3
0.35	13.0	13.1	13.2	13.3	13.5	13.9	14.4	14.8	15.6
0.40	13.1	13.1	13.2	13.3	13.5	13.8	14.3	14.7	15.4
0.45	12.7	12.8	12.9	13.0	13.1	13.4	13.7	14.1	14.6
0.50	12.2	12.2	12.3	12.4	12.5	12.6	12.9	13.1	13.4
0.55	11.7	11.7	11.7	11.8	11.8	11.9	12.1	12.2	12.3
0.60	11.7	11.7	11.7	11.7	11.8	11.9	12.1	12.2	12.4
0.65	12.9	12.9	13.0	13.1	13.3	13.5	13.9	14.3	14.9
0.70	15.9	16.0	16.1	16.4	16.7	17.3	18.1	18.9	20.3
0.75	20.8	20.9	21.2	21.6	22.2	23.1	24.6	25.9	28.1
0.80	27.3	27.6	28.0	28.6	29.4	30.9	33.0	34.9	38.2
0.85	35.5	35.8	36.3	37.2	38.4	40.4	43.3	45.9	50.4
0.90	45.2	45.6	46.3	47.4	48.9	51.6	55.3	58.8	64.6
0.95	56.4	57.0	57.9	59.2	61.2	64.5	69.3	73.6	80.9
1.00	69.3	70.0	71.0	72.7	75.1	79.2	85.1	90.5	99.5

$n$		$\rho_{51}(k, n)$							
$k$	$\infty$	40	25	18	14	11	9	8	7
0.00	15.2	15.2	15.4	15.6	16.0	16.6	17.5	18.4	20.0
0.05	14.9	15.0	15.2	15.4	15.7	16.3	17.1	18.0	19.6
0.10	14.3	14.4	14.5	14.7	15.0	15.5	16.2	16.9	18.2
0.15	13.5	13.5	13.6	13.8	14.0	14.4	14.9	15.5	16.4
0.20	12.7	12.7	12.8	12.9	13.1	13.4	13.8	14.2	14.8
0.25	12.3	12.4	12.5	12.6	12.7	13.0	13.5	13.8	14.5
0.30	12.7	12.8	12.9	13.1	13.3	13.7	14.4	15.0	16.2
0.35	14.0	14.1	14.2	14.4	14.8	15.5	16.5	17.5	19.4
0.40	15.7	15.8	16.0	16.4	16.9	17.7	19.1	20.5	23.1
0.45	17.6	17.7	18.0	18.4	19.0	20.0	21.6	23.3	26.4
0.50	19.2	19.3	19.6	20.0	20.6	21.7	23.4	25.1	28.4
0.55	20.1	20.2	20.5	20.9	21.4	22.4	24.0	25.5	28.3
0.60	20.1	20.2	20.4	20.7	21.1	21.9	22.9	24.0	25.9
0.65	19.4	19.5	19.6	19.7	19.9	20.3	20.7	21.1	21.7
0.70	18.9	18.9	19.0	19.0	19.2	19.3	19.6	19.9	20.6
0.75	21.0	21.1	21.3	21.6	22.2	23.2	25.0	27.2	31.9
0.80	28.5	28.9	29.5	30.5	32.0	35.0	39.9	45.2	56.0
0.85	42.3	43.0	44.2	46.2	49.1	54.6	63.4	72.6	90.8
0.90	62.2	63.4	65.3	68.4	73.1	81.5	95.1	109	136
0.95	88.4	90.1	92.9	97.4	104	116	136	155	194
1.00	121.2	123.6	127.4	134	143	159	186	212	264

TABLE 1 (Continued)

$n$		$\rho_{22}(k, n)$								
$k$	$n$	$\infty$	40	25	18	14	11	9	8	7
all		13.4	13.4	13.5	13.5	13.6	13.7	13.8	14.0	14.1
$n$		$\rho_{32}(k, n)$								
$k$	$n$	$\infty$	40	25	18	14	11	9	8	7
0.00		13.4	13.4	13.5	13.5	13.6	13.7	13.8	14.0	14.1
0.05		14.0	14.0	14.1	14.1	14.2	14.3	14.5	14.7	14.9
0.10		15.6	15.6	15.7	15.8	15.9	16.1	16.3	16.6	16.9
0.15		17.9	18.0	18.1	18.2	18.4	18.6	19.0	19.3	19.8
0.20		20.8	20.9	21.0	21.1	21.3	21.7	22.2	22.6	23.3
0.25		24.0	24.0	24.2	24.4	24.7	25.1	25.8	26.3	27.1
0.30		27.3	27.4	27.6	27.8	28.2	28.7	29.5	30.2	31.2
0.35		30.9	31.0	31.2	31.4	31.9	32.5	33.4	34.2	35.4
0.40		34.5	34.6	34.8	35.2	35.6	36.4	37.4	38.3	39.7
0.45		38.1	38.3	38.6	38.9	39.5	40.3	41.5	42.5	44.0
0.50		41.9	42.1	42.3	42.8	43.3	44.3	45.6	46.7	48.4
0.55		45.7	45.9	46.2	46.6	47.3	48.3	49.8	51.0	52.9
0.60		49.5	49.7	50.0	50.5	51.2	52.4	53.9	55.3	57.3
0.65		53.3	53.5	53.9	54.4	55.2	56.4	58.1	59.6	61.8
0.70		57.2	57.4	57.8	58.4	59.2	60.5	62.4	63.9	66.4
0.75		61.0	61.3	61.7	62.3	63.2	64.7	66.6	68.3	70.9
0.80		64.9	65.2	65.6	66.3	67.2	68.8	70.9	72.7	75.4
0.85		68.8	69.1	69.5	70.3	71.3	72.9	75.1	77.0	80.0
0.90		72.7	73.0	73.5	74.3	75.3	77.1	79.4	81.4	84.5
0.95		76.6	76.9	77.4	78.2	79.4	81.2	83.7	85.8	89.1
1.00		80.5	80.8	81.4	82.2	83.4	85.4	88.0	90.2	93.7
$n$		$\rho_{42}(k, n)$								
$k$	$n$	$\infty$	40	25	18	14	11	9	8	7
0.00		47.0	47.3	47.7	48.5	49.5	51.3	53.9	56.2	60.0
0.05		46.4	46.7	47.1	47.8	48.9	50.6	53.1	55.4	59.1
0.10		44.7	44.9	45.4	46.0	47.0	48.6	50.9	53.0	56.4
0.15		41.9	42.2	42.5	43.1	44.0	45.4	47.4	49.2	52.2
0.20		38.5	38.7	39.0	39.5	40.1	41.3	42.9	44.3	46.7
0.25		34.8	35.0	35.2	35.6	36.1	36.9	38.0	39.1	40.7
0.30		32.0	32.1	32.3	32.5	32.9	33.5	34.3	34.9	36.0
0.35		31.5	31.6	31.8	32.0	32.4	33.0	33.8	34.5	35.5
0.40		34.9	35.1	35.3	35.7	36.3	37.2	38.6	39.8	41.8
0.45		42.5	42.8	43.2	43.9	44.9	46.5	48.9	51.0	54.6
0.50		53.8	54.2	54.9	55.9	57.4	59.9	63.5	66.8	72.3
0.55		67.9	68.5	69.4	70.9	72.9	76.4	81.4	85.9	93.5
0.60		84.4	85.2	86.4	88.2	90.9	95.4	102	108	118
0.65		103	104	105	108	111	117	125	132	145
0.70		123	125	126	129	133	140	150	159	174
0.75		146	147	149	153	157	166	177	188	206
0.80		170	171	174	178	183	193	207	219	240
0.85		195	197	200	204	211	222	238	253	277
0.90		222	224	228	233	241	253	272	288	316
0.95		251	254	257	263	272	286	307	326	357
1.00		282	284	289	295	305	321	344	365	401

TABLE 1 (Continued)

$\begin{matrix} n \\ \backslash \\ k \end{matrix}$		$\rho_{52}(k, n)$								
		$\infty$	40	25	18	14	11	9	8	7
	0.00	47.0	47.3	47.7	48.5	49.5	51.3	53.9	56.2	60.0
	0.05	49.5	49.8	50.4	51.3	52.6	54.7	57.9	61.0	66.3
	0.10	56.0	56.5	57.2	58.5	60.2	63.3	68.1	72.7	81.3
	0.15	64.3	64.9	65.9	67.5	69.8	74.0	80.4	86.7	98.7
	0.20	72.4	73.1	74.3	76.2	79.1	84.1	91.8	99.6	114
	0.25	78.9	79.7	81.0	83.1	86.3	91.8	100	109	125
	0.30	82.7	83.6	84.9	87.1	90.2	95.9	105	113	130
	0.35	83.3	84.1	85.3	87.3	90.3	95.5	103	111	126
	0.40	80.4	81.1	82.1	83.8	86.2	90.4	97	103	115
	0.45	74.8	75.3	76.0	77.1	78.8	81.7	86	90	96
	0.50	69.2	69.5	70.0	70.8	71.8	73.6	76	78	82
	0.55	70.2	70.6	71.3	72.4	74.0	76.8	81	86	96
	0.60	86.1	87.0	88.6	91.2	95.1	102	114	126	151
	0.65	119	121	124	129	137	150	172	194	238
	0.70	168	171	176	184	195	216	249	283	350
	0.75	231	235	242	253	270	299	346	394	487
	0.80	307	313	322	337	359	398	461	524	648
	0.85	397	404	416	435	463	514	595	677	835
	0.90	500	510	524	548	584	648	749	852	1050
	0.95	618	630	648	678	721	800	925	1050	1294
	1.00	751	765	787	823	876	971	1122	1274	1568

$\begin{matrix} n \\ \backslash \\ k \end{matrix}$		$\rho_{33}(k, n)$								
		$\infty$	40	25	18	14	11	9	8	7
	all	52.9	53.1	53.5	54.1	54.9	56.2	57.9	59.4	61.7

$\begin{matrix} n \\ \backslash \\ k \end{matrix}$		$\rho_{43}(k, n)$								
		$\infty$	40	25	18	14	11	9	8	7
	0.00	52.9	53.1	53.5	54.1	54.9	56.2	57.9	59.4	61.7
	0.05	56.9	57.2	57.7	58.4	59.4	61.0	63.4	65.4	68.6
	0.10	67.6	68.0	68.7	69.7	71.3	73.8	77.4	80.6	85.9
	0.15	82.3	82.9	83.8	85.4	87.5	91.1	96.3	101	109
	0.20	99.3	100	101	103	106	111	118	124	134
	0.25	118	119	120	123	126	132	141	149	162
	0.30	137	138	140	143	147	154	165	174	190
	0.35	156	158	160	163	168	177	189	200	218
	0.40	176	178	180	184	190	199	213	226	247
	0.45	196	198	201	205	212	223	238	252	276
	0.50	217	219	222	227	234	246	263	279	305
	0.55	237	239	243	248	256	269	288	305	334
	0.60	257	260	264	270	278	293	313	332	364
	0.65	278	281	285	291	300	316	339	359	393
	0.70	299	301	306	313	323	340	364	386	423
	0.75	319	322	327	335	345	363	389	413	452
	0.80	340	343	348	356	368	387	415	440	482
	0.85	361	364	370	378	390	411	440	467	511
	0.90	382	385	391	400	413	434	466	494	541
	0.95	402	406	412	422	435	458	491	521	571
	1.00	423	427	434	444	458	482	516	548	600

TABLE 1 (Continued)

$k \backslash n$		$\rho_{53}(k, n)$								
$k$	$n$	$\infty$	40	25	18	14	11	9	8	7
0.00		238	241	246	254	265	285	317	349	409
0.05		234	237	242	249	260	280	311	341	400
0.10		222	224	229	236	246	264	291	319	372
0.15		203	205	208	214	223	238	261	284	327
0.20		179	180	183	188	194	205	223	239	270
0.25		155	157	159	162	166	174	185	195	213
0.30		144	145	146	149	153	159	168	177	191
0.35		158	160	162	166	172	183	199	216	247
0.40		204	206	211	218	229	248	279	310	369
0.45		274	278	285	296	313	343	391	439	532
0.50		362	368	378	394	417	460	526	594	724
0.55		465	473	486	507	538	593	681	770	941
0.60		580	590	607	633	672	743	853	965	1181
0.65		708	720	740	772	820	906	1042	1179	1440
0.70		846	861	885	924	981	1084	1247	1410	1730
0.75		996	1013	1042	1087	1155	1276	1470	1660	2030
0.80		1160	1180	1210	1260	1340	1480	1700	1930	2360
0.85		1330	1350	1390	1450	1540	1700	1960	2210	2710
0.90		1510	1540	1580	1650	1750	1930	2220	2520	3080
0.95		1700	1730	1780	1860	1970	2180	2510	2840	3470
1.00		1900	1940	1990	2080	2210	2440	2810	3170	3890

$n$		$\rho_{44}(k, n)$							
$k$	$\infty$	40	25	18	14	11	9	8	7
all	210	212	215	220	227	239	257	272	299

$k \backslash n$		$\rho_{54}(k, n)$								
$k$	$n$	$\infty$	40	25	18	14	11	9	8	7
0.00		210	212	215	220	227	239	257	272	299
0.05		235	237	241	248	257	274	299	323	366
0.10		296	300	307	317	332	359	400	441	519
0.15		377	383	392	407	428	467	527	588	703
0.20		468	475	487	506	535	586	666	746	900
0.25		563	572	587	611	646	710	809	909	1103
0.30		661	672	690	719	761	837	956	1076	1310
0.35		761	774	794	828	877	966	1105	1245	1520
0.40		862	876	900	938	995	1096	1255	1410	1720
0.45		963	980	1006	1049	1113	1226	1410	1590	1930
0.50		1066	1084	1113	1161	1232	1360	1560	1760	2140
0.60		1270	1290	1330	1390	1470	1620	1860	2100	2560
0.70		1480	1500	1540	1610	1710	1890	2160	2440	2990
0.80		1680	1710	1760	1840	1950	2150	2470	2790	3410
0.90		1890	1930	1980	2060	2190	2420	2780	3140	3830
1.00		2100	2140	2200	2290	2430	2680	3080	3480	4260

$k \backslash n$		$\rho_{55}(k, n)$								
$k$	$n$	$\infty$	40	25	18	14	11	9	8	7
all		836	850	874	912	968	1069	1228	1388	1698



than 2 per cent., but may rise to a maximum of about 5 per cent. for the smaller values of  $n$  near the last minimum of the  $\rho(k, n)$  curve where the functions vary rapidly with  $k$ .

In the region of extrapolation the function  $\rho_{pj}(k, n)$  may be written in the form

$$\rho_{pj}(k, n) = \rho_{pj}(k) \cdot \varphi_p(n),$$

where

$$\rho_{pj}(k) = \rho_{pj}(k, \infty)$$

and

$$\varphi_p(n) = D_p^{\frac{1}{2}}.$$

The values  $\rho_{pj}(k)$  are given in Table 2 for  $p$  up to 5. The values  $\varphi_p(n)$  are given in Table 2 of the previous paper. The error introduced by the splitting of  $\rho$  into two factors is less than 1 per cent. except for  $\rho_{51}$  and  $\rho_{52}$  when the value of  $k$  is close to unity. For  $\rho_{51}$  near  $k=1$  the error is less than 1 per cent. for  $n \geq 18$ , and never exceeds 5 per cent. except in the single case  $n=7$ . For  $\rho_{52}$  near  $k=1$ , the error is less than 1 per cent. for  $n \geq 11$  and never exceeds 3 per cent.

#### IV. VARIATION OF THE ESTIMATED ERROR WITH $k$

As might be expected, the variations of the estimated errors of the polynomial coefficients with  $k$  are similar to the variations of the estimated errors of the fitted values. In particular, the estimated error  $r[c_{pj}]$  has  $p-j$

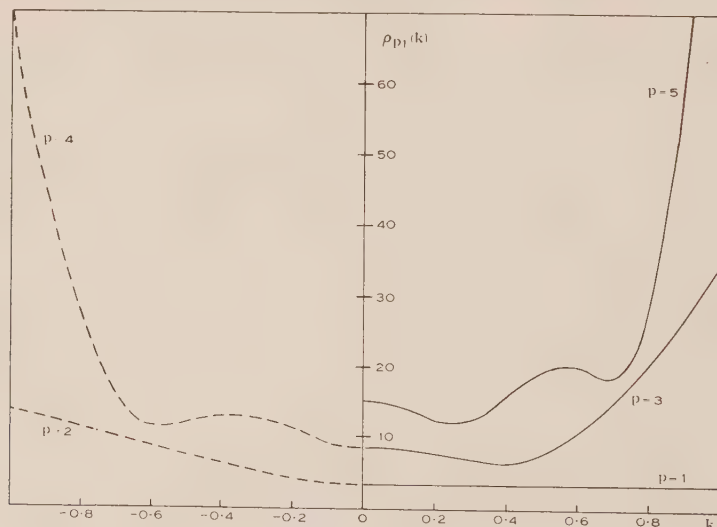


Fig. 1

minima and  $p-j-1$  maxima, all lying in the region of interpolation and symmetrically situated about the point  $k=0$ . Beyond the last minimum the error rises steeply and in the region of extrapolation approaches the form  $r[c_{pj}] = \text{const.} \cdot k^{p-j}$ .

The variation of the estimated error for large  $n$  is shown by the graphs of  $\rho_{p1}(k)$  and  $\rho_{p2}(k)$  in Figures 1 and 2 respectively.

TABLE 2

$k$	$\rho_{21}(k)$	$\rho_{31}(k)$	$\rho_{41}(k)$	$\rho_{51}(k)$
1.00	13.9	34.6	69.3	121
1.05	14.5	38.6	83.7	161
1.10	15.2	42.9	99.9	210
1.15	15.8	47.3	118	266
1.20	16.5	51.9	138	333
1.25	17.1	56.7	159	410
1.30	17.8	61.7	183	499
1.35	18.4	67.0	209	600
1.40	19.1	72.4	236	715
1.45	19.8	78.0	267	844
1.50	20.4	83.9	299	989
1.6	21.7	96.2	371	1330
1.7	23.1	109.2	453	1750
1.8	24.4	123	545	2260
1.9	25.7	138	649	2860
2.0	27.1	153	765	3580
2.1	28.4	169	894	4420
2.2	29.7	187	1036	5390
2.3	31.1	204	1190	6510
2.4	32.4	223	1360	7790
2.5	33.7	242	1550	9260
2.6	35.1	263	1750	10910
2.7	36.4	284	1970	12800
2.8	37.7	306	2200	14900
2.9	39.1	328	2450	17200
3.0	40.4	352	2720	19800
$k > 3$	$13.42k$	$39.7k^2$	$105k^3$	$261k^4$

$k$	$\rho_{32}(k)$	$\rho_{42}(k)$	$\rho_{52}(k)$	$\rho_{43}(k)$	$\rho_{53}(k)$	$\rho_{54}(k)$
1.00	80.5	282	751	423	1900	2100
1.05	84.4	314	900	444	2120	2200
1.10	88.3	348	1066	465	2340	2310
1.15	92.3	383	1249	486	2580	2410
1.20	96.2	420	1450	507	2820	2520
1.3	104.1	498	1910	549	3340	2720
1.4	111.9	583	2450	590	3910	2930
1.5	119.8	674	3080	632	4510	3140
1.6	128	772	3800	674	5160	3350
1.7	136	876	4620	716	5850	3560
1.8	143	986	5550	758	6580	3770
1.9	151	1103	6600	800	7350	3980
2.0	159	1225	7760	842	8170	4180
2.1	167	1350	9050	884	9030	4390
2.2	175	1490	10460	926	9920	4600
2.3	183	1630	12020	967	10860	4810
2.4	191	1780	13700	1009	11800	5020
2.5	199	1930	15600	1051	12900	5230
2.6	207	2090	17600	1093	13900	5440
2.7	215	2260	19800	1135	15000	5650
2.8	223	2430	22100	1177	16200	5850
2.9	231	2610	24600	1219	17400	6060
3.0	238	2800	27300	1261	18600	6270
$k > 3$	$79.4k$	$315k^2$	$1045k^3$	$420k$	$2090k^2$	$2090k$

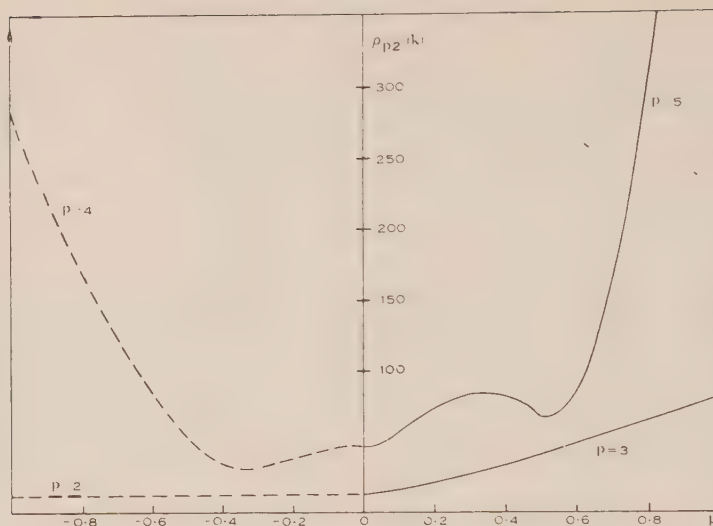


Fig. 2

## V. THE USE OF TABLES

### 1. Estimated Errors of the Polynomial Coefficients

$r[c_{pj}]$  denotes the estimated standard error or the estimated probable error of the coefficient of  $x^j$  in the least-squares polynomial of degree  $p$ . The "physical" variable  $x$  is related to the "statistical" variable  $\varepsilon$  by the equation  $x=h(\varepsilon-m)$ .

The procedure for finding  $r[c_{pj}]$  may be summarized as follows :

- (i) Evaluate  $k=2m/n$  to three decimals,  $n$  being the number of observations.
- (ii) (a) If  $|k| < 1$ , determine the value  $\rho_{pj}(k, n')$  by interpolating between the listed values  $k$  for the listed value  $n'$  closest to  $n$ . Estimate  $\rho_{pj}(k, n)$  by examining the variation with  $n$  in this part of the table.  
 (b) If  $|k| > 1$ , determine  $\rho_{pj}(k)$  by interpolation in Table 2, and  $\varphi_p(n)$  from the table given in the previous paper.
- (iii) (a) If  $|k| < 1$ ,  $r[c_{pj}] = \rho_{pj}(k, n) \cdot r_p / n^{j+\frac{1}{2}} \cdot h^j$ ,  
 (b) If  $|k| > 1$ ,  $r[c_{pj}] = \rho_{pj}(k) \cdot \varphi_p(n) \cdot r_p / n^{j+\frac{1}{2}} \cdot h^j$ ,

where  $r_p$  is the estimated standard error or the estimated probable error of an observation.

### 2. Estimated Errors of the Differential Coefficients

The differential coefficient of order  $j$  with respect to  $x$  at the point on the least-squares curve corresponding to a given value  $m$  of  $\varepsilon$  is  $j!c_{pj}$ , where  $c_{pj}$  is the corresponding power-series coefficient of the polynomial when the origin is transferred to the point  $\varepsilon=m$ . Thus if  $r[c_{pj}]$  is evaluated for  $k=2m/n$  as above, the estimated error of the differential coefficient is  $j!r[c_{pj}]$ .

*Example.*—Consider Birge's Illustrative Problem(2), in which a fourth-degree polynomial was fitted to seven observations. Thus  $n=7$ ,  $p=4$ ;  $r_p$

had the value  $7.25 \times 10^{-2}$ . The variable was changed to  $x = h(\varepsilon - m)$ , where  $m=3$ ,  $h=5$ . Therefore  $k=6/7=0.857$ . Hence we have Table 3.

TABLE 3

$j$	$\rho_{4j}$	$r[c_{4j}]$
1	52.4	$\frac{5.24 \times 10 \times 7.25 \times 10^{-2}}{7 \times 5 \times 2.65} = 4.09 \times 10^{-2}$
2	282	$\frac{2.82 \times 10^2 \times 7.25 \times 10^{-2}}{4.9 \times 10 \times 2.5 \times 10 \times 2.65} = 6.30 \times 10^{-3}$
3	515	$\frac{5.15 \times 10^2 \times 7.25 \times 10^{-2}}{3.43 \times 10^2 \times 1.25 \times 10^2 \times 2.65} = 3.29 \times 10^{-4}$
4	299	$\frac{2.99 \times 10^2 \times 7.25 \times 10^{-2}}{2.40 \times 10^3 \times 6.25 \times 10^2 \times 2.65} = 5.45 \times 10^{-6}$

These figures are practically identical with the values calculated by Birge. Values of  $h^j$ ,  $n^j$ , and  $n^{\frac{1}{2}}$  were taken from Barlow's Tables of Squares etc. All calculations (including interpolation between  $k=0.85$  and  $k=0.90$ ) were done by slide rule.

## VI. REFERENCES

- (1) GUEST, P. G.—*Aust. J. Sci. Res. A* **3**: 173 (1950).
- (2) BIRGE, R. T.—*Rev. Mod. Phys.* **19**: 298 (1947).
- (3) KERAWALA, S. M.—*Indian J. Phys.* **15**: 241 (1941).
- (4) FISHER, R. A. and YATES, F.—“Statistical Tables for Biological, Agricultural, and Medical Research.” (Oliver and Boyd: Edinburgh, 1948.)



# EQUIVALENT PATH AND ABSORPTION FOR ELECTROMAGNETIC RADIATION IN THE SOLAR CORONA

By J. C. JAEGER\* and K. C. WESTFOLD†

[Manuscript received March 28, 1950]

## Summary

Calculations of the trajectories, equivalent path, and absorption of rays, in the frequency range 20–100 Mc/s., in the solar corona have been made, neglecting possible magnetic fields and assuming spherical symmetry.

Interpreting the double-humped burst of solar noise as the superposition of a direct and an echo signal, inferences are made as to the height in the corona and location on the solar disk of its source.

## I. INTRODUCTION

Many observations have recently been made of the electromagnetic radiation emitted by the sun on frequencies of 20 Mc/s. and above. It is now well established that, because of the steady increase of electron density inwards through the corona and chromosphere, waves of frequencies less than about 200 Mc/s. can occur only in the corona while waves of higher frequencies may be present in the chromosphere.

With the development of experiment and theory, quantitative information about the behaviour of rays in the sun's atmosphere, similar to that now available for the ionosphere, will become increasingly necessary: in particular it is needed now for the study of the "unpolarized bursts" of solar noise described by Payne-Scott(1) on frequencies in the range of 20–100 Mc/s. In the present paper numerical information is given about the paths, equivalent paths, and absorption of rays on frequencies of 20, 60, and 100 Mc/s. on certain simplifying assumptions, notably, spherical symmetry, and neglect of magnetic fields. The work has been confined to this frequency range since observations of bursts of solar noise are practically confined to it, and also the available data are more reliable for the corona. Some cross plots interpolating the results enable information to be read off for other frequencies within this range.

As remarked above we neglect the magnetic field because of the difficulty of considering it and because of the uncertainty of its amount. Since the first application of the results will be to bursts of radiation which are not circularly polarized this is probably justifiable: in other cases it may lead to results which are correct in order of magnitude.

We also assume that the electron density in the corona is spherically symmetrical and for its value use Allen's revision(2) of Baumbach's data, namely

$$N = 1.42 \times 10^8 \rho^{-6} (1 + 1.68 \rho^{-10}), \quad \rho \geq 1, \dots\dots\dots (1)$$

\* University of Tasmania, Hobart.

† Division of Radiophysics, University Grounds, Sydney.

where  $N$  is the electron density (electrons/cc.; units are c.g.s. throughout) at a distance  $r$  cm. from the centre of the sun, and  $\rho$  is this distance expressed in units of  $R_1$ , the radius of the base\* of the corona, that is

$$r = \rho R_1, \quad \dots \dots \dots (2)$$

where  $R_1 = 7.05 \times 10^{10}$  cm.

The distribution (1) is fairly well established and is all that is needed for the calculation of path and equivalent path. For the calculation of absorption the collision frequency  $\nu$  is needed, and this is less certain. The value

$$\nu = 42NT^{-3/2}, \quad \dots \dots \dots (3)$$

due to Smerd and Westfold(3), will be used, taking for the kinetic temperature,  $T$ , Alfvén's(4) value

$$T = 1.63 \times 10^6 \frac{\rho^{-1}(1 + 0.69\rho^{-10})}{(1 + 1.68\rho^{-10})}. \quad \dots \dots \dots (4)$$

The numerical factor 42 in (3) is an average of a function of density and temperature which does not vary by more than a few units over the whole corona.

In Sections II, III, and IV, respectively, the path, equivalent path, and absorption of various rays are calculated under these assumptions. In Section V some results for a single inverse power law for  $N$  are given and discussed. In Section VI an application to the location of the sources of bursts of unpolarized radiation is given.

## II. THE TRAJECTORIES

Let  $r, \theta$  be the polar coordinates referred to the centre of the sun of a point on the ray whose asymptote is at distance  $p$  from the axis  $OX$  (Fig. 1). The



Fig. 1.—The coordinates specifying a trajectory.

equation of the path is derived from Snell's law

$$\mu r \sin i = \text{constant}, \quad \dots \dots \dots (5)$$

where  $\mu$  is the refractive index of the medium and  $i$  the angle of incidence of the ray at the point  $(r, \theta)$ . Since  $\mu \rightarrow 1$  and  $r \sin i \rightarrow p$  as  $r \rightarrow \infty$  we may write (5) in the form

$$\mu r \sin i = p. \quad \dots \dots \dots (6)$$

\* This value of course is not well known and Allen's results were expressed in units of the base of the chromosphere. In the present paper  $R_1$  only appears as a scale factor, but in any extension to higher frequencies different laws of variation of electron density have to be used for the chromosphere and corona, and it is convenient to have  $\rho > 1$  for the corona and  $\rho < 1$  for the chromosphere.

As in Section I it is convenient to express all lengths in units of  $R_1$ , the radius of the base of the corona, so we write

$$r = \rho R_1, \quad p = a R_1, \quad \dots \dots \dots (7)$$

and (6) becomes

$$\mu \rho \sin i = a. \quad \dots \dots \dots (8)$$

The parameter  $a$  introduced in (7) is the fundamental one which will be used to specify the rays. It follows from (8) that

$$\frac{\mu^2 \rho^2}{a^2} = \operatorname{cosec}^2 i = 1 + \frac{1}{\rho^2} \left( \frac{d\rho}{d\theta} \right)^2,$$

so that the equation of the path is

$$\theta = a \int_{\rho}^{\infty} \frac{d\rho}{\rho \sqrt{(\mu^2 \rho^2 - a^2)}}, \quad \dots \dots \dots (9)$$

the limits being chosen so that  $\theta \rightarrow 0$  as  $\rho \rightarrow \infty$ .

The refractive index  $\mu$  is known by the usual Lorentz theory to be

$$\mu^2 = 1 - \frac{Ne^2}{\pi m f^2} \times 10^{-12},$$

where  $f$  is the frequency of the wave in Mc/s. and  $e$  and  $m$  are the charge and mass of the electron. Using numerical values for these and the value (1) of  $N$  gives

$$\mu^2 = 1 - \frac{11500}{f^2} \rho^{-6} (1 + 1.68 \rho^{-10}). \quad \dots \dots \dots (10)$$

Clearly in both (9) and (10) it is more convenient to use the parameter  $1/\rho$ , so we write

$$\eta = 1/\rho. \quad \dots \dots \dots (11)$$

In terms of this (9) becomes

$$\theta = a \int_0^{1/\rho} \frac{d\eta}{\Delta^{\frac{1}{2}}(a, \eta)}, \quad \dots \dots \dots (12)$$

where, for shortness, we write

$$\Delta(a, \eta) = 1 - (11500/f^2) \eta^6 (1 + 1.68 \eta^{10}) - a^2 \eta^2 \quad \dots \dots \dots (13)$$

The distance of closest approach  $R_1 \rho_a$  for any given values of  $a$  and  $f$  is given by  $\rho_a = 1/\eta_a$ , where  $\eta_a$  is the root of  $\Delta(a, \eta) = 0$ , that is, of

$$1 - \frac{11500}{f^2} \eta_a^6 (1 + 1.68 \eta_a^{10}) - a^2 \eta_a^2 = 0. \quad \dots \dots \dots (14)$$

$\rho_a$  and the value  $\theta_a$  derived from (12) with  $\rho = \rho_a$  are plotted in Figure 2A against  $a$  for frequencies of 20, 60, and 100 Mc/s. The polar coordinates of the point of closest approach of the ray are  $(\rho_a R_1, \theta_a)$ . The integrals for  $\theta$ , and all those considered subsequently, are evaluated by numerical integration. In Figure 2B,  $\rho_a$  and  $\theta_a$  are cross plotted against frequency for  $a = 0, 0.25, \dots, 1.5$ .

In Figures 3 and 4 polar plots of the paths for the frequencies of 60 and 100 Mc/s. and various values of  $a$  are given. It appears that the paths only

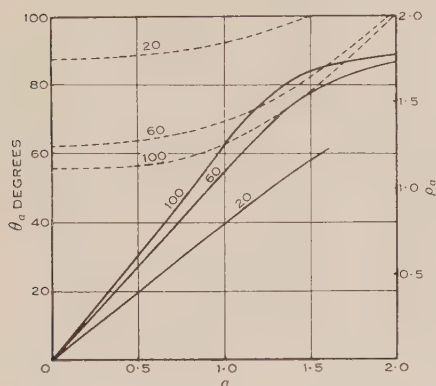


Fig. 2A

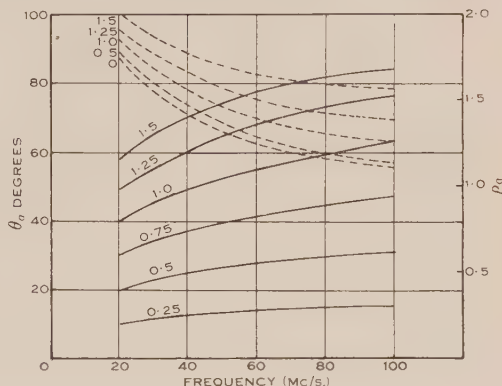


Fig. 2B

Figs. 2A and 2B.—Polar coordinates  $\rho_a$  (dotted lines) and  $\theta_a$  (full lines) of the point of closest approach. The numbers on the curves are, in Figure 2A, frequencies in Mc/s. and, in Figure 2B, the values of  $a$ .

deviate from their asymptotes over a fairly small range of values of  $\rho$  so that they can easily be sketched roughly from a knowledge of their asymptotes and the position of the point of closest approach given in Figures 2A and 2B. Finally

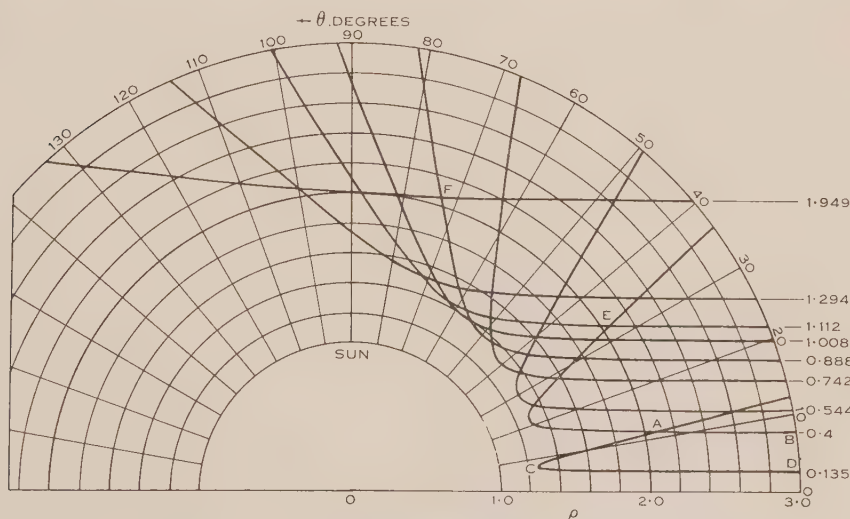


Fig. 3.—Trajectories for 60 Mc/s. The numbers on the curves are the values of  $a$ .

it should be remarked that the values of  $a$  used in Figure 3 and subsequent figures are chosen to give simple values of  $\eta_a$  (such as 0.80, 0.79, 0.78, 0.76, ...) for convenience in the numerical work.



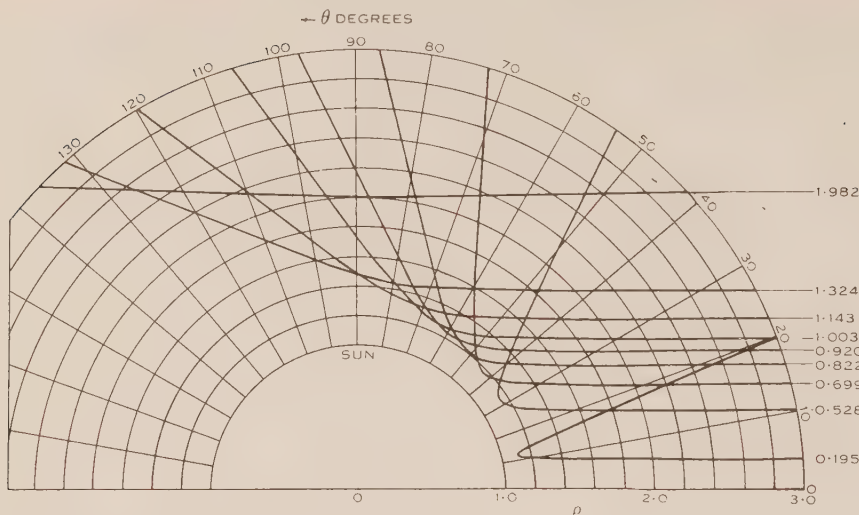


Fig. 4.—Trajectories for 100 Mc/s. The numbers on the curves are the values of  $a$ .

### III. THE EQUIVALENT PATH

If  $ds$  is the element of arc of the path, the equivalent path\* is defined by

$$\int \frac{ds}{\mu} = R_1 \int \frac{d\rho \sec i}{\mu} = R_1 \int \frac{\rho d\rho}{\sqrt{(\mu^2 \rho^2 - a^2)}}, \dots \dots \dots (15)$$

taken over the portion of the trajectory in question. Putting  $\eta = 1/\rho$  and using the notation (13) this becomes

$$R_1 \int \frac{d\eta}{\eta^2 \Delta^{\frac{1}{2}}(a, \eta)}. \dots \dots \dots (16)$$

The integral, of course, is divergent as  $\eta \rightarrow 0$  so we evaluate

$$P(a, \rho) = \int_0^{1/\rho} \frac{d\eta}{\eta^2} \left\{ \frac{1}{\Delta^{\frac{1}{2}}(a, \eta)} - 1 \right\} \dots \dots \dots (17)$$

$R_1 P(a, \rho)$  is thus the excess of equivalent path along the chosen trajectory from infinity to the point distant  $\rho R_1$  radially from the origin over the distance along the axis from infinity to a point at the same distance  $\rho R_1$  from the origin.

In Figure 5,  $P(a, \rho)$  is plotted against  $\rho$  for various values of  $a$  for a frequency of 60 Mc/s. and in Figure 6 a similar set of curves is given for 100 Mc/s.

### IV. THE ABSORPTION INTEGRALS

We shall calculate

$$F(a, \rho) = \int \kappa ds, \dots \dots \dots (18)$$

taken along a trajectory from infinity to any value of  $\rho$ , where by the usual Lorentz theory the absorption coefficient  $\kappa$  is

$$\kappa = \frac{\nu}{2c} \frac{1 - \mu^2}{\mu}, \dots \dots \dots (19)$$

\* The equivalent path along portion of a trajectory is the time taken for a signal to traverse it multiplied by  $c$ , the velocity of light.

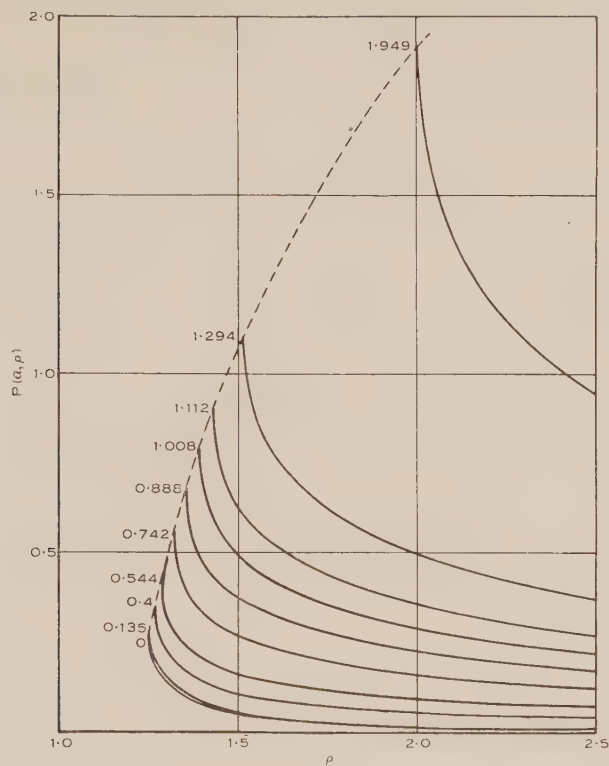


Fig. 5.—Equivalent path integrals for 60 Mc/s. The numbers on the curves are the values of  $a$ .

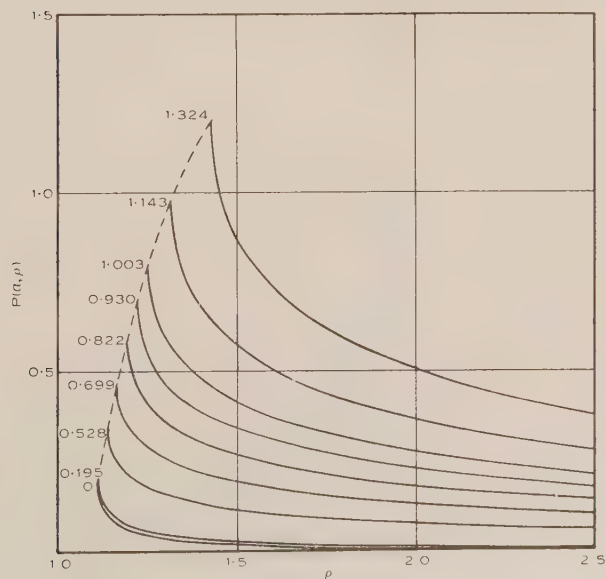


Fig. 6.—Equivalent path integrals for 100 Mc/s. The numbers on the curves are the values of  $a$ .

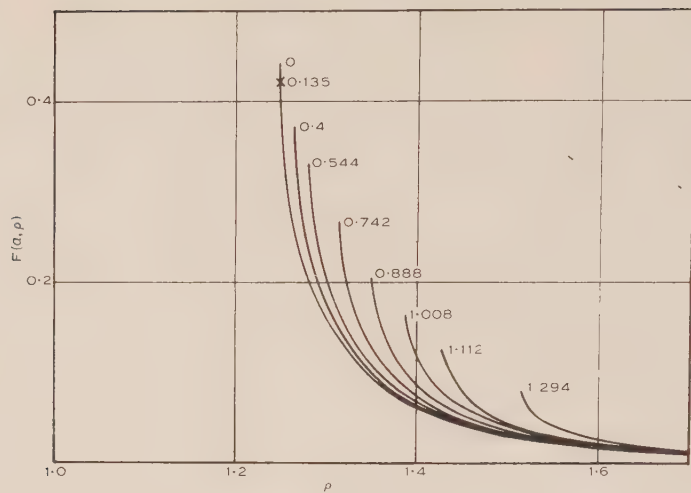


Fig. 7.—Absorption integrals for 60 Mc/s. The numbers on the curves are the values of  $a$ .

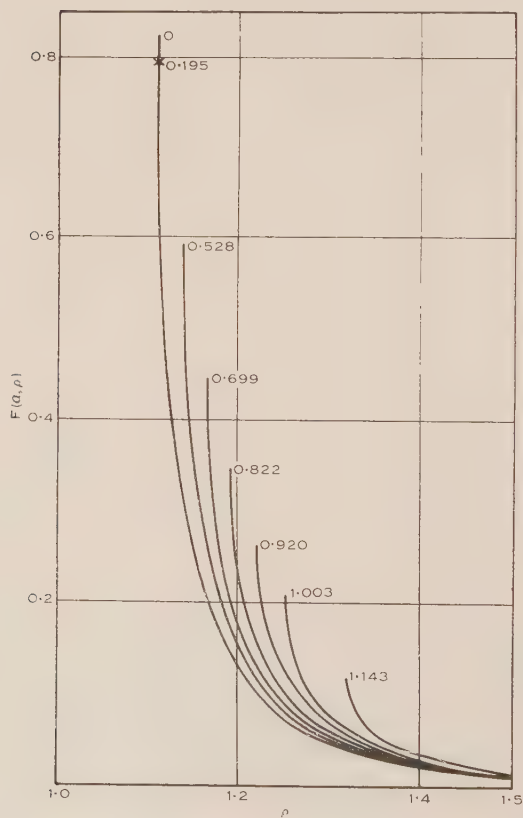


Fig. 8.—Absorption integrals for 100 Mc/s. The numbers on the curves are the values of  $a$ .

$c$  being the velocity of light. The optical depth along a trajectory, as usually defined, is  $2F(a, \rho)$ .

The essential part of the integrand consists of the term  $ds/\mu$  which may be written

$$\frac{ds}{\mu} = -\frac{R_1 d\eta}{\eta^2 \Delta^{\frac{1}{2}}(a, \eta)}, \quad \dots \dots \dots (20)$$

as in Section III.

Inserting this and the values (3) and (4) of the collision frequency  $\nu$ , also using numerical values of the constants  $R_1$ ,  $c$ , etc. in the integral (18), we get finally

$$F(a, \rho) = \frac{38700}{f^2} \int_0^{1/\rho} \frac{\eta^{17/2} (1 + 1.68\eta^{10})^{7/2} d\eta}{(1 + 0.69\eta^{10})^{3/2} \Delta^{\frac{1}{2}}(a, \eta)}. \quad \dots \dots \dots (21)$$

In Figures 7 and 8, values of this integral for frequencies of 60 and 100 Mc/s., respectively, are plotted against  $\rho$  for various values of  $a$ . In Figure 9A the values of  $F(a, \rho_a)$ , the absorption integrals from infinity to the point of closest

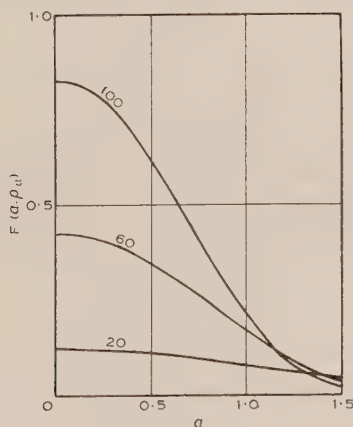


Fig. 9A

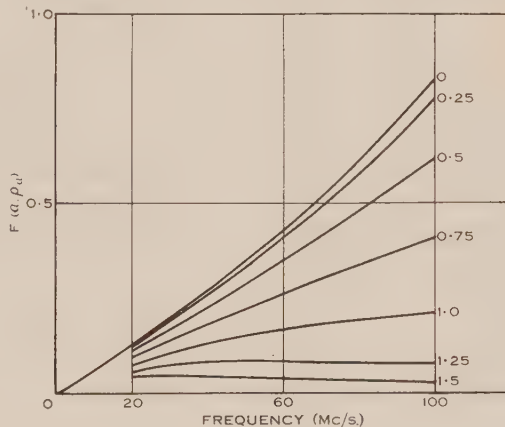


Fig. 9B

Figs. 9A and 9B.—Absorption integrals  $F(a, \rho_a)$  from infinity up to the point of closest approach. The numbers on the curves are, in Figure 9A, the frequencies in Mc/s. and, in Figure 9B, the values of  $a$ .

approach, are plotted against  $a$  for frequencies of 20, 60, and 100 Mc/s. and in Figure 9B,  $F(a, \rho_a)$  is cross plotted against frequency for  $a=0, 0.25, \dots, 1.5$ .

## V. THE INVERSE POWER LAW OF ELECTRON DENSITY

It is well established that the electron density over the whole of the corona cannot be represented by a single inverse power law, and the combination (1) of two such terms seems to be the simplest adequate representation. At the longer distances such as those involved in studying 20 Mc/s. radiation, the effect of the second term in (1) is negligible and we are left with an inverse sixth power. For this law some accurate results are obtainable which are of interest, both because of their application to 20 Mc/s. radiation, and because they enable the effect of some of the assumptions made above to be studied.



Thus we consider the electron density

$$N = N_0 \rho^{-n}, \dots\dots\dots (22)$$

and to obtain simple closed results we must restrict ourselves to the axial ray  $a=0$  and to integrals from infinity to the point of closest approach. These integrals can be expressed in terms of gamma functions; the results are given below.

(i) *Absorption from infinity to the point of closest approach of the axial ray for electron density (22) and kinetic temperatures given by*

$$T = T_0 / \rho. \dots\dots\dots (23)$$

The distribution (23) is obtained from (4) by neglecting terms in  $\rho^{10}$ .

The absorption integral in this case is

$$1.08 \times 10^{6-15/n} (81)^{5/2n} T_0^{-3/2} N_0^{5/2n} f^{2-5/n} \frac{\Gamma(2-5/2n)}{n \Gamma(5/2-5/2n)}. \dots\dots (24)$$

For the present case in which  $n=6$ ,  $N_0=1.42 \times 10^8$ ,  $T_0=1.63 \times 10^6$ , this becomes

$$0.00365 f^{7/6}. \dots\dots\dots (25)$$

For  $f=20$  this gives 0.120 in place of 0.123 as in Figure 9A.

(ii) *Absorption from infinity to the point of closest approach of the axial ray for electron density (22) and constant kinetic temperature  $T_1$ .*

In this case we get for the absorption integral

$$1.08 \times 10^{6-6/n} (81)^{1/n} N_0^{1/n} \frac{\Gamma(2-1/n) f^{2-2/n} T_1^{-3/2}}{n \Gamma(5/2-1/n)}. \dots\dots (26)$$

Taking  $n=6$ ,  $N_0=1.42 \times 10^8$ , and  $T_1=10^6$ , this becomes

$$6.75 \times 10^{-4} f^{5/3}, \dots\dots\dots (27)$$

which gives 0.100 for  $f=20$ .

This result illustrates the differences to be expected from different models. We have used the formula (4) for temperature since it has a theoretical justification and gives a reasonable decrease at large distances; the alternative is to assume a constant kinetic temperature, the value  $10^6$  degrees being usually taken.

(iii) *The excess of equivalent path for the axial ray over distance traversed from infinity inwards to the point of closest approach.*

This is found to be

$$R_1 (81 \times 10^{-6})^{1/n} N_0^{1/n} f^{-2/n} \left\{ 1 - \frac{\Gamma(\frac{1}{2}) \Gamma\left(1 - \frac{1}{n}\right)}{\Gamma\left(\frac{1}{2} - \frac{1}{n}\right)} \right\}. \dots\dots (28)$$

If  $n=6$  and  $N_0=1.42 \times 10^8$  this becomes

$$1.203 f^{-1/3} R_1. \dots\dots\dots (29)$$

For  $f=20$  this gives the value  $0.44 R_1$  which agrees with the value  $0.44 R_1$  obtained from (17). The values for  $f=60$  and 100 obtained from (17) are lower than those given by the simple formula (29).

It appears that for the inverse power law the excess of equivalent path over distance decreases with increasing frequency. It may be remarked that this behaviour is different to that of other common distributions: for an exponential distribution the excess is independent of frequency, and for a Chapman distribution it increases with increasing frequency.

## VI. APPLICATION TO UNPOLARIZED BURSTS OF SOLAR NOISE

It has been observed by Payne-Scott(1) that the majority of "unpolarized bursts" of solar noise on frequencies of 60–100 Mc/s. show a pronounced "double-humped" form as in Figure 10.

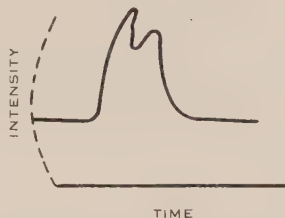


Fig. 10.—A typical "unpolarized burst".

A possible explanation of this is that the burst is caused by a localized disturbance at some point of the corona at which radiation of the frequency in question can be propagated, and that radiation from the disturbance reaches the observer by either of the two possible trajectories. For example, radiation on 60 Mc/s. from a disturbance at *A* (Fig. 3) can reach the observer by either of the paths *AB* or *ACD*. The "echo" signal which follows the path *ACD* will be more attenuated and will arrive later than the "direct" signal along *AB* by amounts which can be calculated from Figures 5 and 7.

If records such as Figure 10 are interpreted in this way, the time delay between the direct and echo signals and the ratio of their intensities can be determined. Most observations have been made on 60 Mc/s., and for this frequency the time delays range from about 2 to 10 seconds, while the ratio of the intensities of the direct and echo signals varies from about 2:1 to 8:1. A time-delay of 2 seconds, corresponding to a difference in equivalent path between the two rays of the order of  $R_1$ , implies in general by Figure 5 that the source of the disturbance is well out in the corona, say  $\rho > 1.6$ , and this in turn implies by Figure 7 that the absorption in the "direct" trajectory is negligible and that the ratio of the intensities of the two signals is determined almost entirely by attenuation in the echo path. Thus the ratio of the intensities in the direct and echo signals gives the parameter  $a$  which specifies the echo path. For example, on 60 Mc/s. it follows from Figures 7 or 9A that the ratios of the intensities for  $a=0, 0.135, 0.4, 0.888$ , respectively, are 5.9:1, 5.5:1, 4.5:1, and 2.3:1. When the echo path has been determined, the position of the source on it can be found from the time delay, using Figure 5.

In general, for a source at a fixed distance from the centre of the sun, both the ratio of the intensities and the time delay decrease as the source moves

across the disk. To examine this point quantitatively, consider sources of 60 Mc/s. radiation at the points  $A$ ,  $E$ , and  $F$  of Figure 3, all approximately at a distance of  $2.06 R_1$  from the centre of the sun. For the source at  $A$  the time delay is 5 seconds and the ratio of the intensities  $5.5:1$ , very little different from the values 5.0 seconds and  $5.9:1$  at the centre of the disk. For the source at  $E$ , a little off the disk, the values are 4.5 seconds and  $4.5:1$ , while for the source at  $F$ , at an angle of  $73^\circ$  with the direction to the observer, the values are 2.4 seconds and  $2.3:1$ . As the angle increases to  $90^\circ$  or beyond they decrease still further. Since bursts in which the components have a time delay of less than two seconds and much the same intensity would be recorded as "single-humped", it follows that bursts of this type may, on the present hypothesis, be due either to a relatively distant source at an angle of nearly  $90^\circ$  from the observer's direction, or, of course, to nearer sources than those considered above.

Similar remarks apply to 100 Mc/s. radiation except that, because of the larger attenuation, sources near the centre of the disk will probably be recorded as "single-humped".

If these conclusions are substantiated in the main when the positions of sources can be determined by interferometry, it will be possible to check the constants in the formulae (1), (3), and (4) by comparing observed and calculated quantities for a source in a known position.

#### VII. ACKNOWLEDGMENT

We are indebted to Miss M. E. Clarke of the University of Tasmania for assistance with the numerical work.

#### VIII. REFERENCES

- (1) PAYNE-SCOTT, RUBY.—*Aust. J. Sci. Res. A* **2**: 214-27 (1949).
- (2) ALLEN, C. W.—*Mon. Not. R. Astr. Soc.* **107**: 426-32 (1947).
- (3) SMERD, S. F., and WESTFOLD, K. C.—*Phil. Mag.* (7) **40**: 831-48 (1949).
- (4) ALFVÉN, H.—*Ark. Mat. Astr. Fys.* **27**: No. 25 (1941).

# OBSERVATIONS OF THE SPECTRUM OF HIGH-INTENSITY SOLAR RADIATION AT METRE WAVELENGTHS

## I. THE APPARATUS AND SPECTRAL TYPES OF SOLAR BURST OBSERVED

By J. P. WILD\* and L. L. MCCREADY\*

[*Manuscript received March 10, 1950*]

### *Summary*

An apparatus for recording the dynamic spectrum of high-intensity solar radiation (in particular the sudden bursts) in the frequency range 70–130 Mc/s. is described. The spectra are displayed on a cathode-ray tube at intervals of about one-third of a second.

Solar bursts observed with the apparatus were found to have widely different spectra. However, analysis of a number of bursts indicated the common occurrence of three distinct spectral types. These types are described and illustrated by samples. One type, of narrow bandwidth, was exhibited by short-lived bursts that occur in large numbers during periods of high intensity ("noise storms"); these bursts are presumed to be circularly polarized and associated with sunspots. A second type, characterized by a slow drift of spectral features towards the lower frequencies, was exhibited by sporadic "outbursts" associated with solar flares. Other sporadic bursts had diverse spectra, but some of them conformed to a third spectral type in which the frequency of maximum intensity drifts rapidly towards the lower frequencies.

The result that outbursts seem to exhibit a distinct type of spectrum is considered to provide a possible means of recognizing these phenomena with certainty.

## I. INTRODUCTION

During recent years a considerable amount of evidence concerning the high-intensity radio-frequency radiation emitted by the sun has been assembled through observations made with receiving devices tuned to fixed radio frequencies.† In particular it has been established that sudden short-lived increases or "bursts" of radio-frequency radiation are emitted by the sun from time to time. Investigation of the frequency spectrum of these phenomena has hitherto depended upon simultaneous observations at two or three frequencies. Spectral information obtained in this way is necessarily incomplete and its detailed interpretation speculative. Reliable spectral data clearly require observations which extend over a continuous frequency range.

This paper is the first of a series describing some observations that were made over a continuous frequency range between about 70 and 130 Mc/s., i.e. between wavelengths of about 4.3 and 2.3 metres. The apparatus and method of observation are described in Sections II and III. Some typical results indicating the different types of spectra observed in solar bursts are given in

\* Division of Radiophysics, C.S.I.R.O., University Grounds, Sydney.

† A survey of the early work has recently been given by Pawsey(1).



Section IV. The relation of these types to the kinds of burst described by observers using conventional (single-frequency) equipment is discussed in Section V. A detailed analysis of the results will be given in subsequent papers.

## II. APPARATUS

### (a) *The Method*

To analyse the spectrum of swiftly changing phenomena an instrument is required which registers the spectrum sufficiently often to allow the structure of the most rapid of the variations to be recorded. These requirements were satisfied, in the instrument to be described, by receiving radiation with a broad-band aerial connected to a receiver which was rapidly tuned over the frequency range. The tuning arrangements of the receiver were mechanically driven at a speed such that the frequency sweep occupied 0.07 second, and repeated itself about three times per second. The spectrum was displayed on a cathode-ray tube in the form of a graph of receiver output (vertical) *versus* frequency (horizontal). The display was photographed. Thence, from a knowledge of the frequency-dependent characteristics of the aerial and the receiver, the records were converted into graphs relating the intensity of the radiation at the aerial to the frequency of the radiation.

### (b) *The Aerial*

A rhombic aerial was used to receive the radiation. This type of aerial was chosen because it is able to receive radiation along its principal axis (i.e. its axis of maximum acceptance) over a broad band of frequencies with little variation of gain. For solar observations the principal axis was maintained in the direction of the sun. The angle subtended by the source at the aerial was small compared with the width of the main lobe of the aerial pattern; hence only those aerial characteristics which affect propagation along the principal axis need be considered.

The intensity of radiation at the aerial due to a source in its principal axis will be referred to in terms of the quantity  $S_f$ , such that  $S_f df$  is the *accepted plane-polarized component* of the power flowing normally through unit area in the frequency interval between  $f$  and  $f+df$ . Throughout this paper the quantity  $S_f$  will be referred to as the "intensity" of the radiation.\*

If  $P_f df$  is the power available at the terminals of the aerial within the same frequency interval, the "effective area",  $A_f$ , of the aerial in the direction of its principal axis is defined by

$$A_f = \frac{P_f}{S_f} \dots\dots\dots (1)$$

It is desirable that the effective area of the aerial be as constant as possible over the frequency range.

The rhombic aerial is shown diagrammatically in Figure 1 in which each side is of length  $l$ , and the angle between the sides at the corners  $X$  is  $2\phi$ . The aerial's characteristic impedance,  $Z$ , was made approximately uniform by

\* For radiation with no preferred plane of polarization the "intensity" defined here is equal to half the total "intensity". Some authors use the latter quantity to express their results.

constructing each side of two wires of equal length, arranged in a plane normal to that of the rhombus, which diverge towards the corners  $X$ . The far end is terminated in a resistive load to minimize reflections at this point.

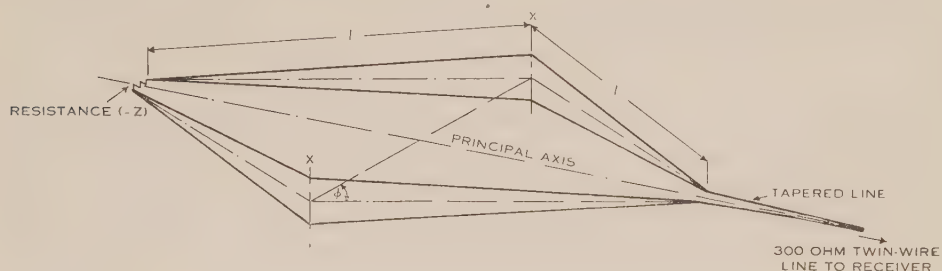


Fig. 1.—The rhombic aerial.

The evaluation of the effective area in terms of the parameters  $l$ ,  $\phi$ , and  $Z$ , and the factors involved in the design of the aerial to satisfy the prescribed frequency range are discussed in Appendix I; the values assigned to  $l$ ,  $\phi$ , and  $Z$  in the actual aerial are also given. The final relation between  $A_f$  and  $f$  is plotted in Figure 2.

The aerial was mounted in such a way that its principal axis could be rotated conveniently about the polar axis at any desired declination.

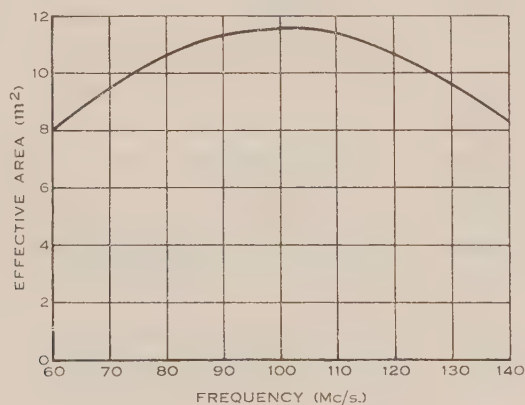


Fig. 2.—The variation with frequency of the effective area of the rhombic aerial used for the measurements described in this paper. The curve is calculated from equation (A6) in Appendix I.

The input impedance of the aerial was approximately a resistance of 670 ohms throughout the frequency range. A tapered line, three metres long, transformed this resistance to one of 300 ohms, whence connection was made with the receiver by means of a two-wire line of characteristic impedance 300 ohms. Considerable care was taken to eliminate, as far as possible, the presence of standing waves in this line, since any mismatch would cause errors in the calibration. Tests carried out at the receiver end of the line showed, however, that a slight mismatch occurred at the lower frequencies of the range. This

mismatch caused small errors in the calibration at these frequencies ; the errors reached a maximum at two or three frequencies spaced at about 7 Mc/s. intervals—the periodicity associated with the length of the feeder—but at no frequency were they as much as two to one in power. The effects of these errors on the final spectra are described in Section III.

### (c) *The Receiver*

The receiver was of conventional superheterodyne design with a single stage of radio-frequency amplification. The local oscillator and the input circuits of the radio-frequency amplifier and the mixer were tuned by “split-stator” condensers on a single shaft which was rotated by a motor at about three revolutions per second. The condensers passed through their complete tuning range as the shaft was rotated through a right angle ; the instrument was arranged to operate during only one such quadrant, referred to subsequently as the *operative quadrant*. By using split-stator condensers, the need for making electrical contacts with rotating shafts was avoided.

The bandwidth of the radio-frequency circuits was made sufficiently broad to allow effective tracking of the tuned circuits over the band : the “resolving power” of the instrument was then determined by the bandwidth of the intermediate-frequency amplifier which was fixed at about 300 kc/s. The bandwidth of the radio-frequency circuits was, however, limited by two factors : the demand for high sensitivity (and, therefore, of efficient tuned circuits) ; and the necessity of suppressing the unwanted “image” frequency. The necessary variables in the radio-frequency circuits were provided to allow precise alignment at any two frequencies ; such alignment of the broad-band circuits was found sufficient to give a fairly constant gain and noise factor over the band.

The output of the second detector was amplified for display on a cathode-ray tube. The time constant of the video circuits was of the order of  $10^{-4}$  seconds, this value being a compromise chosen to satisfy two conflicting requirements :

- (1) Reduction of random fluctuations about the mean level of the “noise” output (requiring a large time constant).
- (2) Response, without distortion, to abrupt discontinuities in the spectrum of the radiation despite the rapidity of the frequency sweep (requiring a short time constant).

### (d) *The Display*

The amplified receiver output was applied to the Y-plates of a cathode-ray tube.

The frequency sweep (see Fig. 3) was controlled by a variable condenser  $C$ , coupled to the receiver tuning shaft. An oscillatory potential of constant amplitude was applied across a high resistance,  $R$ , in series with a quarter-wave transmission line terminated by the condenser  $C$ . Now the capacity of  $C$  was directly proportional to the angular displacement of the tuning shaft ; so also was the impedance, and hence the potential, across the terminals  $AB$  of the transmission line. The oscillatory potential at  $AB$  was rectified, and, after amplification, fed to the X-plates of the cathode-ray tube, thus providing a

suitable frequency sweep. With these arrangements the picture appearing on the screen represented a graph of receiver output (vertically) *versus* frequency (horizontally).

The final appearance of the display is illustrated in Plate 1. The receiver output appeared on the scan only during the operative quadrant. (During the other three quadrants an external potential was switched across the Y-plates to deflect the spot completely off the screen.) The zero level of the receiver output was displayed at each end of the sweep; this was arranged by switching a large negative bias onto the grid of one of the valves of the intermediate frequency amplifier for short periods at the beginning and end of each operative quadrant. In both cases the switching was accomplished by photoelectric

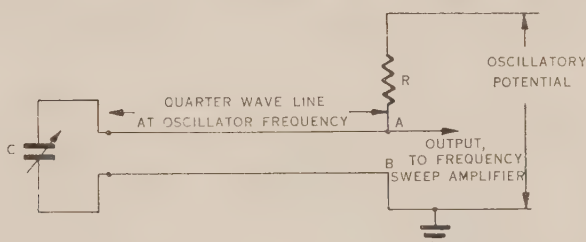


Fig. 3.—The fundamental circuit of the frequency sweep. The condenser  $C$  is ganged to the receiver tuning shaft. The output from terminals  $AB$  is subsequently rectified and amplified, and applied to the X-plates of a cathode-ray tube.

means, thereby eliminating mechanical contacts: photoelectric cells were exposed to small lamps through slotted disks which rotated on the receiver-tuning shaft; the cell currents were subsequently amplified and fed to the Y-plates and the intermediate-frequency amplifier respectively.

Time was recorded by photographing a revolution counter (situated next to the display) operated by the tuning shaft. The number of revolutions was logged at hourly intervals determined by broadcast time signals.

#### (e) Calibration

The ultimate purpose of the instrument was the determination of the intensity,  $S_f$ , of the received radiation at all frequencies in the range. Now in accordance with equation (1), the determination of  $S_f$  involves the measurement of three quantities: the effective area,  $A_f$ , of the aerial; the power,  $P_f$ , available at the terminals of the aerial; and the frequency  $f$ . The evaluation of  $A_f$  has already been discussed in Subsection II (b).

In order to determine the available power  $P_f$ , the photographed records of receiver output *versus* frequency were calibrated by comparing the recorded amplitudes with those given by replacing the aerial by a source of similar impedance and known noise output. A noise generator containing a temperature-limited diode was used for the calibrating source; the impedance of the source was made to simulate that of the aerial by inserting a dissipative network between the noise generator and the receiver. Plate 1d shows a typical



picture of the display when the aerial was replaced by the noise generator; the dotted line shows the level of the scan when the current through the diode was zero—in that case the level was determined merely by the noise inherent in the receiver together with a small contribution from the thermal noise produced in the source resistance. It can be seen from the photograph that the sensitivity of the receiver was fairly uniform over the frequency range.

The scan was calibrated in frequency by injecting signals of known frequency into the receiver before each set of observations was taken. The resulting "pips" on the scan were photographed and provided a set of frequency markers.

By combining the three sets of calibration data (i.e. those determining  $A_f$ ,  $P_f$ , and  $f$ ) and by suitable interpolation, a comprehensive intensity-frequency grid was constructed. Projection of the final photographed records upon this grid enabled intensities and frequencies to be read off directly.

### III. THE METHOD OF OBSERVATION AND PRESENTATION OF RESULTS

The observations described in this paper were taken at Penrith, N.S.W., between February and June 1949. The total time of observation amounted to 264 hours; throughout this time the display was watched by an observer who operated a ciné-camera, capable of photographing each scan appearing on the display whenever phenomena of interest appeared. In this way records of many solar bursts were obtained.

The photographic record of a single burst comprises a number of pictures, each of the type illustrated in Plate 1*a-c*, separated by intervals of about one-third of a second. If no marked irregularities in the spectral profile occur during these intervals, a record of this kind gives a fairly complete description of the dynamic spectrum of a burst within the prescribed frequency limits.

The information is conveniently displayed by time-frequency diagrams in which the intensity (as defined in Subsection II (*b*)) is specified by contours (see Plate 2). Successive contours have been chosen to represent a logarithmic scale of intensity, the interval between any two representing a two-to-one change in intensity. In each diagram, time is measured from an arbitrary zero point. The beginnings of the bursts were often not recorded; in general the method of recording introduced a time lag between the onset of a burst and the starting of the recording camera.

Throughout the period of observation, the cosmic-noise level remained appreciably below the lowest contour level (arbitrarily chosen at  $5 \times 10^{-21}$  watts m.<sup>-2</sup> (c/s.)<sup>-1</sup>) and has not been subtracted from the total intensities in this analysis.

Interfering signals from radio transmitters appeared on the display from time to time; they were identified either by listening through a loudspeaker or by the characteristic suddenness with which they appeared and disappeared. The effects of such signals were eliminated from the diagrams.

Small systematic errors due to aerial mismatch (see the end of Subsection II (*b*)) occur in the diagrams. These may produce two or three slight spurious "ridges", separated by about 7 Mc/s., parallel to the time axis at the lower frequencies. Fortunately, these effects can be recognized at once; the ridges

at 78 and 85 Mc/s. in Plate 2*c*, for instance, are spurious. Quite generally, however, any peak encompassed by two or more contours can be interpreted as a genuine solar phenomenon.

#### IV. OBSERVATION OF BURSTS OF DISTINCT SPECTRAL TYPES

The dynamic spectra of the observed bursts showed widely different features and were often complicated. A detailed description of the bursts must, therefore, be lengthy and will be undertaken in later papers. The analysis showed, however, that the spectra of many of the observed bursts conformed to one of three distinct types. The results to be given in this paper are restricted to indicating the salient features of these types.

Typical examples of the three spectral types are shown in Plate 2 *a*, *b*, and *c*. (The types correspond to instantaneous spectra shown in Plate 1*a*, *b*, and *c* respectively.) We shall consider them in turn:

*Type I.*—Plate 2*a* is typical of bursts which occurred in hundreds during restricted periods, called “noise storms” by Allen(2). Such periods usually lasted for hours, or even days, and were often marked by a high, slowly-varying background continuum above which the bursts appeared. Bursts of this spectral type rarely occurred as isolated phenomena. The type is characterized by its narrow spectrum, a few megacycles per second wide. (The burst at 112 Mc/s. in Plate 2*a* appears broader towards the finish because of confusion with smaller bursts at 115 and 121 Mc/s.) The lifetime of the bursts is sometimes less than one second and sometimes as long as 20 seconds.\* The frequency of maximum intensity remains approximately constant throughout the lifetime, although this intensity may fluctuate rapidly. The bursts seem to occur at all frequencies in the band.

These “storm bursts” constituted the great majority of those observed. The other types of burst, discussed below, occurred sporadically, either singly or in small groups.

*Type II.*—Plate 2*b* illustrates a type of burst of comparatively rare occurrence. In this type, the spectrum at any instant near the start shows a distinct “cut-off” in frequency, little or no radiation being received at frequencies immediately below a critical frequency. The critical frequency varies with time and, on the average, *drifts gradually towards the lower frequencies* at a rate of the order of  $\frac{1}{4}$  (Mc/s.) sec.<sup>-1</sup>. The bursts last for several minutes (only the first part is shown in the illustration). The start of the burst illustrated coincided approximately with the start of a solar flare of importance two.† Of five such bursts observed, three coincided with a large solar flare or sudden short-wave communication fade-out.

\* Bandwidths and lifetimes are specified in this paper by the intervals between points of quarter-maximum intensity.

† Observed at the Commonwealth Observatory, Canberra, A.C.T.

About 30 other sporadic bursts, or small groups of them, were observed. None coincided with a reported solar flare or sudden fade-out. They tended to be of short duration (about 3 to 30 seconds) and broad bandwidth (not less than about 15 Mc/s.) but were otherwise diverse in their spectral characteristics. Rather less than half, however, conformed to a third distinct spectral type :

*Type III.*—In this type, illustrated in Plate 2c, the frequency of maximum intensity drifts rapidly (at a rate of the order of 20 (Mc/s.) sec.<sup>-1</sup>) towards the lower frequencies. The bandwidth of these bursts at any instant is not usually less than about 50 Mc/s. They last for a few seconds.

#### V. THE RELATION BETWEEN SPECTRAL TYPES AND CLASSES OF BURSTS OBSERVED WITH CONVENTIONAL EQUIPMENT

It is of interest to examine the relation between the spectral types discussed above and the types of burst described by observers using single-frequency receiving apparatus.

It has been established by Payne-Scott(3) that high-intensity solar radiation at metre wavelengths can be divided into two broad groups according to the polarization of the radiation, as follows\* :

- (1) Circularly-polarized radiation. This is the "enhanced radiation" which predominates during noise storms. It is closely associated with the meridian passage of large sunspot groups, and is emitted from the vicinity of sunspots. It appears on the single-frequency records as a slowly-varying background level above which short-lived bursts, of the same polarization as the background, occur. The bursts show little or no correspondence at frequencies differing by more than a few megacycles per second. These bursts are clearly the "storm bursts", discussed above, which conform to the spectral type I.
- (2) Radiation which is not circularly polarized. This radiation occurs in sporadic bursts which last for some seconds or, in extreme cases, several minutes. The bursts are often observed simultaneously, or nearly so, at widely different frequencies.

Certain of these sporadic bursts, particularly the larger ones, have been observed to coincide significantly with solar flares(4), and the name "outburst" has been given to them(2). Nevertheless it has not been possible to draw an objective line of demarcation between the outbursts (defined as bursts typically associated with flares) and other sporadic bursts. The observation of the spectrum, however, provides further information which may allow the outbursts to be distinguished by observable properties. For whereas bursts of spectral type II showed a good correlation with flares and sudden fade-outs, no instance of a coincidence was found with other sporadic bursts. This result indicates that *outbursts conform, in some cases at least, to a distinct spectral type and can therefore be recognized at once from a record of the spectrum.* Further evidence,

\* The properties given here are based on the work of various observers ; for references see Payne-Scott(3) and Pawsey(1).

based on a large number of observations, is required to determine whether outbursts *always* conform to the one type.

In view of the slow frequency-drift characteristic of spectral type II, we might expect that, in simple cases such as that illustrated in Plate 2*b*, outbursts would start at markedly different times at different frequencies. An instance was reported by Payne-Scott, Yabsley, and Bolton(5) in which such time delays were observed. It was suggested that the delays were associated with an exciting cause moving outwards through the solar atmosphere. This hypothesis will be examined in the succeeding paper.

Because of the association of outbursts with flares, it seems proper to regard them as a class by themselves. The other sporadic bursts, called "isolated bursts" by Pawsey(1), are usually of considerably shorter duration than the outbursts. The results given in Section IV indicate that these isolated bursts are of diverse spectra, although their bandwidth tends to be broad in all cases. The observation that some of them conform to spectral type III (characterized by a rapid drift in frequency) is consistent with reports(3, 5) that bursts have a tendency to arrive at slightly different times on different frequencies, high frequencies preceding low frequencies.

The characteristics of bursts discussed above are summarized in Table 1. In the subsequent papers we shall consider the different types separately, giving a detailed analysis and statistical results in each case.

TABLE 1  
TENTATIVE CHARACTERISTICS OF BURSTS

Class of Burst	Occurrence	Polarization	Visual Correlation	Dynamic Spectrum			
				Type	Bandwidth	Duration	Remarks
Storm bursts	During noise storms	Circular	Sunspots	I	Narrow	Seconds	
Outbursts	Sporadic	Not circular	Flares	II	Broad	Minutes	Complex. Slow frequency drift towards lower frequencies
Isolated bursts	Sporadic	Not circular	—	Diverse including III	Broad	Seconds	Bursts of type III spectrum show rapid drift towards lower frequencies



## VI. ACKNOWLEDGMENTS

The authors are indebted to Mr. J. D. Murray for the design and construction of the display apparatus and associated equipment, and to Mr. W. C. Rowe for considerable assistance throughout the work described in this paper. The authors also wish to thank Dr. J. L. Pawsey and Mr. S. F. Smerd for valuable advice and criticism in the preparation of this paper.

## VII. REFERENCES

- (1) PAWSEY, J. L.—Solar Radio-frequency Radiation, *J. Instn. Elect. Engrs.* (in press).
- (2) ALLEN, C. W.—*Mon. Not. R. Astr. Soc.* **107**: 386 (1947).
- (3) PAYNE-SCOTT, RUBY.—*Aust. J. Sci. Res. A* **2**: 214 (1949).
- (4) HEY, J. S., PARSONS, S. J., and PHILLIPS, J. W.—*Mon. Not. R. Astr. Soc.* **108**: 354 (1948).
- (5) PAYNE-SCOTT, R., YABSLEY, D. E., and BOLTON, J. G.—*Nature* **160**: 256 (1947).
- (6) BRUCE, E., BECK, A. C., and LOWRY, L. R.—*Proc. Inst. Radio Engrs.* **23**: 24 (1935).
- (7) CHRISTIANSEN, W. N.—*A.W.A. Tech. Rev.* **7**: 361 (1947).

## EXPLANATION OF PLATES 1, 2

## PLATE 1

Photographs of the cathode-ray tube display showing typical examples of instantaneous spectral profiles of the three types of bursts discussed in Section IV: (a) Type I; (b) Type II; (c) Type III. In these cases the dotted lines indicate the "quiet" level of the undisturbed sun. (d) shows the effect of a broad-band calibration signal from a noise generator; here the dotted line indicates the zero-signal level.

*Explanation of display.*—Receiver output is displayed vertically and frequency (ranging from 70 Mc/s. at the left-hand side to 130 Mc/s. at the right-hand side) horizontally. In all cases the zero level of the receiver output (indicated on the display by bright markers at each end) is inclined to the horizontal because the display was photographed by a ciné-camera with a continuously moving film.

*Note.*—The break in the trace occurring slightly left of centre in (a) and (c) is due to an interference signal (92 Mc/s.).

## PLATE 2

Examples of dynamic spectra of bursts illustrating the three spectral types discussed in Section IV.

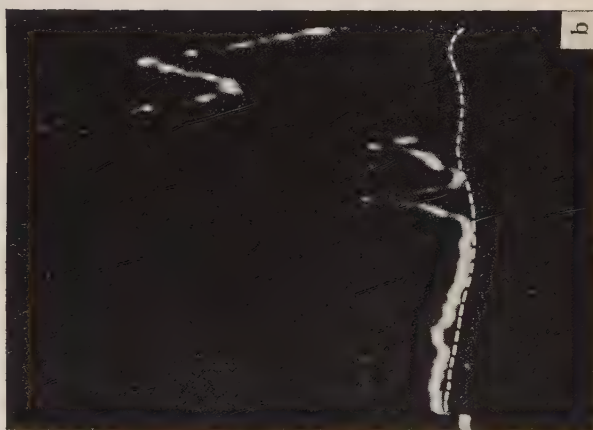
- (a) Type I observed on May 10, 1949, at 01 hr. 46 min. (G.M.T.).
- (b) Type II observed on June 27, 1949, at 02 hr. 40 min.
- (c) Type III observed on May 31, 1949, at 01 hr. 09 min.

## APPENDIX I

*The Effective Area and Design of the Rhombic Aerial*

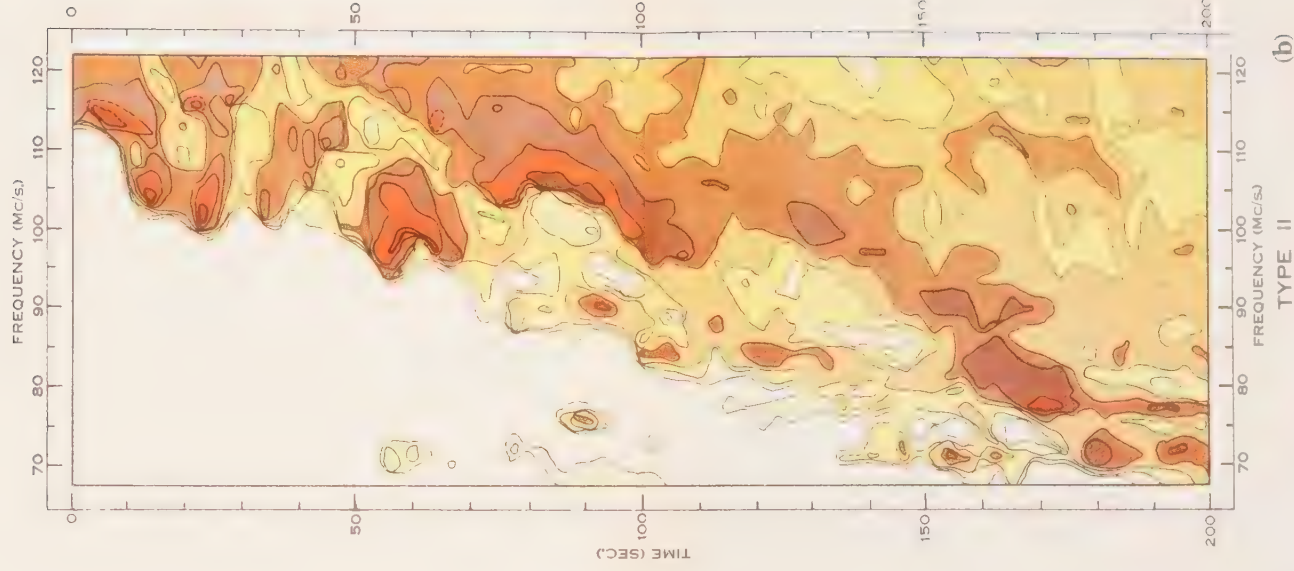
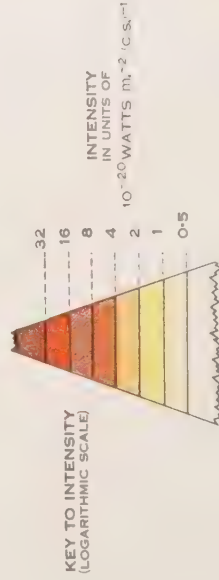
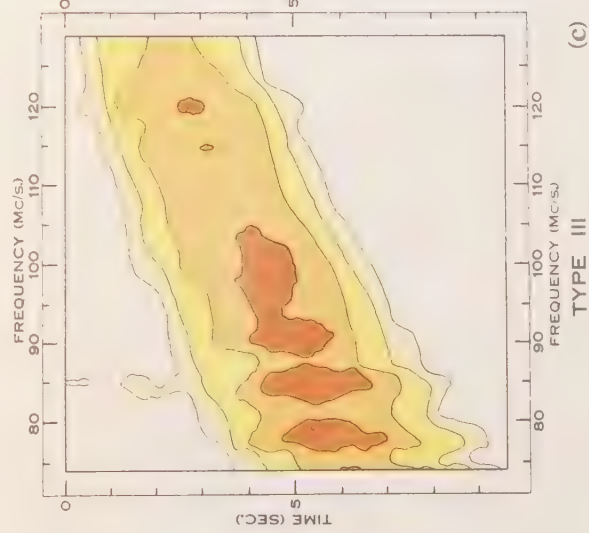
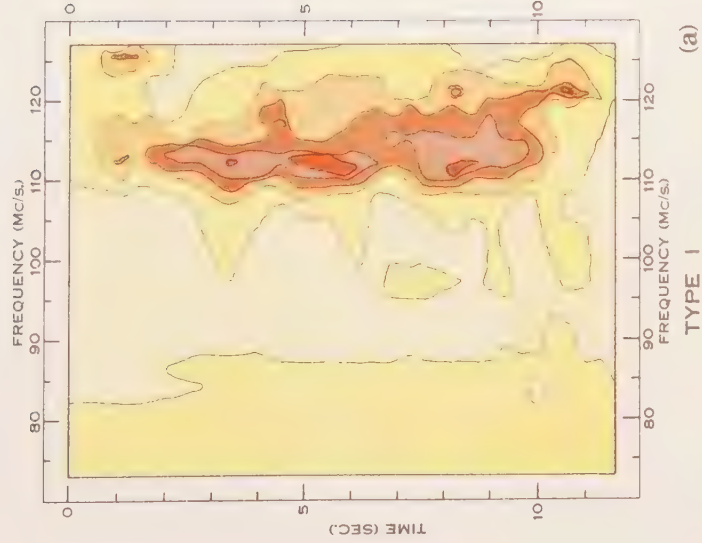
The configuration of a rhombic aerial is shown in Figure 1. We shall calculate the effective area (defined by equation (1)) of such an aerial for radiation received along its principal axis. The effect of ground reflections will be neglected and the current distribution in the wires will initially be assumed to be uniform. Under these conditions the current in a receiver of input impedance  $Z$  is given by the equation of Bruce, Beck, and Lowry(6) as

$$I = \frac{iE\lambda}{2\pi Z} \left( \frac{\cos \varphi}{1 - \sin \varphi} \right) \left\{ 1 - \exp \left[ -\frac{2\pi i l}{\lambda} (1 - \sin \varphi) \right] \right\}^2,$$



WILD AND MCCREADY.—OBSERVATIONS OF THE SPECTRUM OF HIGH-INTENSITY  
SOLAR RADIATION AT METRE WAVELENGTHS. I









where  $I$  is the complex receiver current (r.m.s. value),

$E$  is the field strength at the aerial due to a distant source in the principal axis of the aerial (r.m.s. value),

$\lambda$  is the wavelength of the radiation,

$i$  is the  $\sqrt{-1}$ ,

$\varphi$ ,  $l$ , and  $Z$  are as defined in Subsection II (b) and Figure 1.

Hence the modulus of  $I$  is

$$|I| = \frac{2E\lambda F(\varphi, l/\lambda)}{\pi Z},$$

where

$$F(\varphi, l/\lambda) = \frac{\cos \varphi \sin^2 \left\{ \frac{\pi l}{\lambda} (1 - \sin \varphi) \right\}}{1 - \sin \varphi}. \quad \text{..... (A1)}$$

Now the amount of radiation,  $S(f)$ , which flows along the principal axis through unit area (normal to the axis) per unit time is, by Poynting's theorem, given by

$$S(f) = \frac{E^2}{120\pi},$$

where  $f$  denotes the frequency of the radiation, and  $S(f)$  and  $E$  are expressed in m.k.s. units. Also the total power available at the terminals of the aerial is given by

$$P(f) = |I|^2 Z.$$

Thus the effective area of the aerial is given by

$$A_f = \frac{P(f)}{S(f)} = \frac{120\pi |I|^2 Z}{E^2} = \frac{480\lambda^2 \{F(\varphi, l/\lambda)\}^2}{\pi Z}. \quad \text{..... (A2)}$$

We now consider the choice of the parameters  $l$  and  $\varphi$  in an aerial designed for maximum effective area over a range of wavelengths between  $\lambda_1$  and  $\lambda_2$  (corresponding to frequencies of  $f_1$  and  $f_2$ ) with particular reference to the special case when  $\lambda_2 = 2\lambda_1$ .

For optimum "tuning" of the aerial to the prescribed frequency limits, the parameters are chosen to make the effective areas at the two extreme frequencies equal, the effective area reaching its maximum value between these two frequencies. This condition is evidently satisfied by the appropriate solution of

$$A_{f_1} = A_{f_2}.$$

It follows from (A2) that this condition is satisfied when

$$\lambda_1 F(\varphi, l/\lambda_1) = \lambda_2 F(\varphi, l/\lambda_2). \quad \text{..... (A3)}$$

In the particular case of a 2 : 1 frequency range,  $\lambda_2 = 2\lambda_1$ , the appropriate solution to (A3) is

$$\sin \varphi = 1 - \lambda_1/2l. \quad \text{..... (A4)}$$

Substituting this value in (A1) gives

$$\{F(\varphi, l/\lambda_1)\}^2 = \frac{4l - \lambda_1}{\lambda_1} \sin^4 \left\{ \frac{\lambda_1 \pi}{\lambda} \frac{\pi}{2} \right\},$$

In the present paper some observations of the spectrum of outbursts over a continuous frequency range (nominally 70–130 Mc/s.) are described, and the correlation of the outbursts with solar flares and their geophysical accompaniments is examined. A possible interpretation of the results is given in Section V.

## II. OBSERVATIONS

The apparatus used for recording the spectrum of high-intensity solar radiation in the frequency range 70–130 Mc/s. has been described in Part I. The observations described in the present paper were taken at intermittent times between February and June 1949, and during this period five outbursts (i.e. bursts of Type II) were observed. The “dynamic” spectra (i.e. spectra showing temporal as well as frequency variations of intensity) are presented by means of contour diagrams in which intensity is specified by contours in the time-frequency plane. The scale of the intensity contours is logarithmic. Spectra of four of the five observed outbursts are shown thus in the accompanying coloured plates. (In each outburst illustrated, only the first part has been recorded owing to limitations in the recording apparatus.)

Details of the five outbursts are as follows :

### *Outburst No. 1*

*Time of start.*—June 27, 1949, 02 hr. 40 min. (G.M.T.).

*Other activity on day of occurrence.*—Slight “noise storm” (i.e. period of high level of solar noise with frequent short-lived bursts) persisting all day.

*Dynamic spectrum* (see Plate 1a).—The spectrum at any one instant near the beginning of the outburst showed an *abrupt cut-off in frequency*. Below the cut-off frequency no appreciable radiation was detected. Above the cut-off frequency the radiation was intense, often showing high narrow-band peaks; such a peak commonly occurred close to the cut-off frequency. The location of the cut-off frequency fluctuated with time but on the average *drifted steadily towards the lower frequencies*. About three minutes after the start, the cut-off frequency had drifted to the lowest frequency in the range (68 Mc/s.), after which it was no longer observable. During the subsequent three minutes (not shown in the plate), the spectrum extended over the entire observed range (68–122 Mc/s. in this case), but showed signs of subsiding at the higher frequencies.

### *Outburst No. 2*

*Time of start.*—April 6, 1949, 05 hr. 04 min.

*Other activity on day of occurrence.*—Few short-lived isolated bursts.

*Dynamic spectrum* (see Plate 1b).—Cut-off and frequency drift as in outburst No. 1. For most of the time, the record can be broadly interpreted as three “drifting” peaks, spaced about 10–15 Mc/s. apart. The relative size of the peaks is in ascending order of frequency. Only the first 80 seconds were recorded. The outburst lasted for about 5 minutes.

*Outburst No. 3*

*Time of start.*—June 8, 1949, 02 hr. 17 min.

*Other activity on day of occurrence.*—Small number of “storm” bursts following outburst.

*Dynamic spectrum* (see Plate 2a).—Cut-off and frequency drift, as in outburst No. 1. The intensity at the cut-off edge rises to an intense peak. On the high frequency side of this peak there is evidence of a second and less clearly defined peak. Only the first two minutes were recorded. The outburst lasted for about six minutes, and was observed to subside at the high frequencies first.

*Outburst No. 4*

*Time of start.*—February 14, 1949, 01 hr. 32 min.

*Other activity on day of occurrence.*—Slight noise storm following outburst.

*Dynamic spectrum* (see Plate 2b).—Cut-off and frequency drift, as in outburst No. 1. The spectrum during the first 50 seconds is not typical of other outbursts observed and may be due to the chance occurrence of a short-lived isolated burst. The characteristic outburst starts at about 55 seconds. From this time onwards the spectrum shows two intense bands separated by a well-defined cut-off band. The two intense bands first appeared at about the same time, and subsequently drifted in parallel formation. Recording ceased after 4 minutes.

*Outburst No. 5*

*Time of start.*—February 10, 1949, 02 hr. 38 min.

*Other activity on day of occurrence.*—Slight noise storm following outburst.

*Dynamic spectrum* (not illustrated).—Only the final phase was observed. When first seen, the spectrum spread over the entire range (72–128 Mc/s. in this case). Later, the extreme high-frequency end subsided completely and the spectrum showed an abrupt cut-off with only the *lower* frequencies present. During the subsequent few minutes this cut-off drifted slowly towards the lower frequencies, eventually below 72 Mc/s., the low-frequency limit.

## III. RECURRENT SPECTRAL FEATURES

We may note that in each of the spectra described above, the following features were present:

- (i) The occurrence of “emission” and “quiescent” bands, separated by an abrupt frequency cut-off.
- (ii) Irregular drift of the emission bands towards the lower frequencies. (The characteristic rate of drift is discussed below.)

These may be regarded as the *defining* features of spectral Type II by which the outbursts were recognized. Other features of interest were: the common occurrence of fluctuating narrow-band peaks within the emission bands, especially near the “cut-off” edges; the occurrence in at least two cases of several peaks within the one emission band, each drifting “parallel” with the cut-off frequency; and the occurrence in one case of two distinct emission bands which drifted in parallel formation.

The frequency drift of the four outbursts illustrated is shown quantitatively in Figure 1. Points on the curves correspond to the position of the low-frequency



"cut-off" edge in each case. The curves have been smoothed for short period variations. The starting point of each curve has been adjusted to allow superposition of the curves. The dotted line drawn through the curves represents a constant drift rate of 0.22 Mc/s. per second. The diagram shows that the average drift-rates of the four outbursts did not depart greatly from this value. It also indicates that, on the average, *the drift rate did not vary appreciably with frequency.*

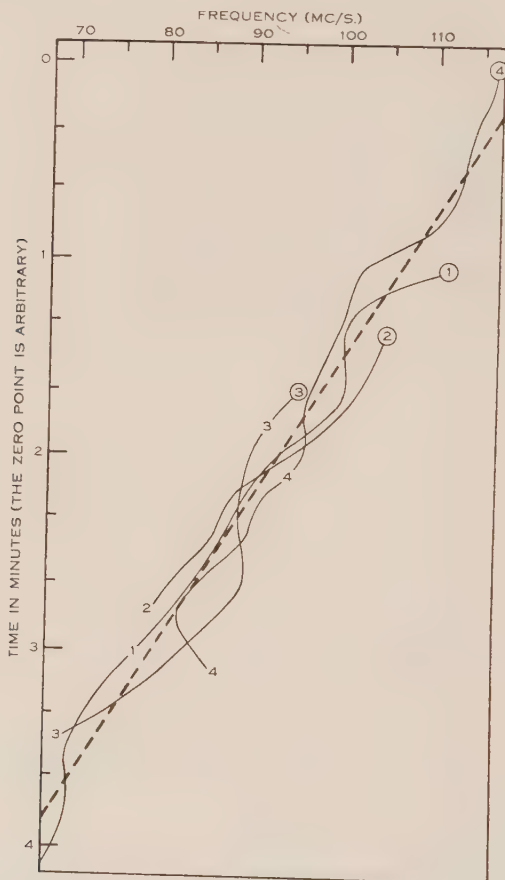


Fig. 1.—The frequency drift of the four outbursts shown in Plates 1 and 2. Smoothed curves of the variation of the frequency of the low-frequency cut-off with time are drawn for each outburst. The four curves have been superposed by adjustment of the starting points (marked 1 for outburst No. 1 etc.). The dotted line represents a constant drift rate of 0.22 Mc/s. per second.

#### IV. THE CORRELATION OF OUTBURSTS AND OTHER SOLAR RADIO-FREQUENCY EMISSIONS WITH FLARES AND THEIR GEOPHYSICAL ACCOMPANIMENTS

In Table 1 is given the correlation of the observed outbursts with solar flares and two of their geophysical accompaniments, viz. sudden communication fade-outs (which characteristically start within a few minutes of the onset of

TABLE I  
CORRELATION OF OBSERVED OUTBURSTS WITH FLARES, FADE-OUTS, AND MAGNETIC STORMS

No.	Outburst		Flare				Fade-out		Magnetic Storm (with Sudden Commencement)
	Date	Time of Start (G.M.T.)	Start*†	Maximum*	Importance	Position on Sun	Start*	Degree	
1	27.vi.49	02 hr. 40 min.	—5 min.	—3 min.	2	8° N. 17° W.	—	—	—
2	6.iv.49	05 hr. 04 min.					—	—	+29.6 hr.
3	8.vi.49	02 hr. 17 min.					—	—	—
4	14.ii.49	01 hr. 32 min.	<—2 min.	+1 min.	2	10° N. 38° E.	—2 min.	Partial	—
5	10.ii.49	02 hr. 38 min.					—8 min.	Complete	+28.2 hr.

\* Times before (—) or after (+) start of outburst.

† No optical observations available.

the flare); and magnetic storms with sudden commencements (which may start about 20–40 hours later).\*† Optical observations were available at the time of two of the outbursts; a flare of importance 2 occurred on both occasions. Three of the observed outbursts coincided with a flare or a fade-out (or both); of the two remaining outbursts, one showed a plausible connexion with a magnetic storm. These data suggest a possible association of outbursts with solar flares and their geophysical accompaniments.

As stated in Part I, the outbursts were the only *sporadic* bursts—i.e. bursts not associated with noise storms—to show any correlation with solar flares. Two instances were, however, recorded during a particularly severe noise storm in which the circularly-polarized background level of enhanced radiation(8) showed an increase to an abnormally high level at the time of a flare; these cases may be examples of another type of disturbance associated with flares and will be discussed further in a subsequent paper dealing with enhanced radiation.

Table 2 gives the correlation of flares reported during the period of observation with outbursts and the abnormal increases in enhanced level. Of nine flares reported, two of importance 2 were accompanied by an outburst, one of importance 2 and one of importance 1 by an increase in enhanced level, and five of importance 1 were unaccompanied. This correlation is consistent with the statistical result of Hey, Parsons, and Phillips(7) that the most intense flares are the most likely to produce radio bursts.

TABLE 2  
CORRELATION OF FLARES REPORTED DURING THE PERIOD OF OBSERVATION  
WITH OUTBURSTS AND INCREASES IN ENHANCED LEVEL

Flares	Coincidences		Uncorrelated
	Outbursts	Increases in Enhanced Level	
Importance 2	2	1	0
Importance 1	0	1	5

Although the number of available observations is small, combination of the data of Tables 1 and 2 indicates a high correlation of outbursts with the larger

\* Correlation with flares and fade-outs is more likely to be significant than correlation with magnetic storms; the link between a given flare and a given magnetic storm is more difficult to establish because of the uncertainty of the delay time between the onsets of the two phenomena.

† The flare and fade-out data given in this paper were provided by the Commonwealth Observatory, Canberra, through the courtesy of Dr. R. v. d. R. Woolley, Commonwealth Astronomer, and Dr. C. W. Allen. Information on magnetic storms is derived from data published in *J. Geophys. Res.* 54: 193, 301 (1949).

solar flares and their geophysical accompaniments, and suggests that outbursts have a definite physical connexion with flares.

### V. INTERPRETATION OF THE OUTBURST SPECTRA

Prior evidence suggesting the occurrence of frequency drift in outbursts has been given by Payne-Scott, Yabsley, and Bolton(4). They reported the observation of a large burst on March 8, 1947, in which time delays of several minutes elapsed between the onset of the burst at 200, 100, and 60 Mc/s. respectively. They suggested that the burst was related to some physical agency moving outwards through the solar atmosphere; this explanation was based on ideas put forward by Martyn(9) that the radiation at any one frequency originates near the level where the coronal electron density reduces the refractive index to zero. We now examine this hypothesis in the light of the evidence given in this paper.

#### (a) *The Interpretation of Observed Outburst Spectra in Terms of the Motion of Localized Disturbances in the Solar Atmosphere*

It is well known that electromagnetic waves can be propagated in an ionized gas only at frequencies higher than the frequency,  $f_0$ , at which the refractive index reduces to zero, given by

$$f_0^2 = \frac{e^2 N}{\pi m}, \quad \dots \dots \dots (1)$$

where  $N$  is the electron density of the medium,

$e$  is the electronic charge (e.s.u.),

$m$  is the electronic mass.

The spectrum due to macroscopic disturbances in such a medium will consequently exhibit a complete cut-off in the range of frequencies below  $f_0$ .\* In a non-uniform medium, such as the solar atmosphere, a cut-off will occur if the disturbance is bounded, the cut-off frequency being determined by equation (1) with  $N$  equal to the lowest electron density at the boundary of the disturbance. If we assume the electron density of the solar atmosphere to decrease continuously outwards, the cut-off frequency will be determined by the electron density at the outermost boundary of the disturbance.

These considerations offer a plausible method of interpreting the observed spectra of outbursts. Such spectra characteristically show an abrupt cut-off at the low-frequency limit of an emission band, and it seems reasonable to identify this cut-off frequency with the frequency  $f_0$  defined above. Hence, if we know the value of  $N$  at different heights in the solar atmosphere, it is possible, by this hypothesis, to derive the height of the outer boundary of the disturbance at any instant from the observed low-frequency cut-off. Assuming values of  $N$

\* The discussion is initially confined to the field-free case. In the presence of magnetic fields, additional frequencies are introduced in accordance with the magneto-ionic theory, and the range of "cut-off" frequencies becomes modified. In the presence of magnetic and electric fields, however, propagation may, according to Bailey(10), take place in the "forbidden" region under certain conditions.



given by Allen's(11) modification of Baumbach's empirical law of coronal electron densities, Figure 2 represents the variation of height with time of the leading edge of the disturbance when the drift rate is that indicated by the dotted line in Figure 1 (drift rate= $0.22 \text{ Mc/s. sec.}^{-1}$ ). We see that observed drift rates correspond to disturbance velocities of the order of 500 km./sec., and that a drift rate which is independent of frequency (suggested as typical by the data of Section III) implies *acceleration outwards through the solar atmosphere*.

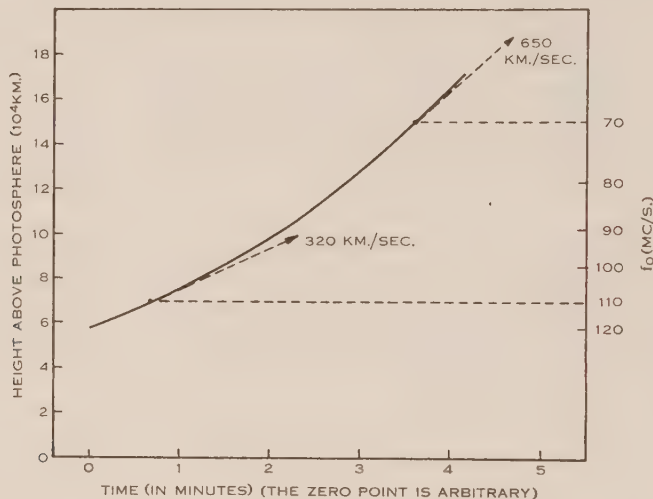


Fig. 2.—The curve indicates the motion in the solar atmosphere of an agency capable of causing an outburst, whose frequency drift corresponds to the dotted line in Figure 1, according to assumptions given in the text.

With regard to the interpretation we have placed on the abrupt cut-off at the low-frequency limit of an emission band, it is noteworthy that a narrow-band peak of high intensity was commonly observed in the immediate vicinity of the cut-off frequency. The occurrence of such peaks is suggestive of some kind of resonance phenomenon taking place near the  $f_0$ -level, as discussed theoretically by Martyn(9) and Jaeger and Westfold(12).

In several of the observed outbursts, the spectrum showed two or three separate peaks or bands at the same time, each drifting at about the same rate. One possible explanation of this behaviour is that a single physical disturbance excites the two or three frequencies at which the refractive index reduces to zero in the presence of a magnetic field(9, 13). On this basis the observed frequency separations would lead to magnetic fields of the order of 10–30 gauss; fields of this magnitude are thought to exist above sunspots at the coronal heights in question(14). However, the different components of the radiation would be expected to show characteristic polarizations, circular or linear. Polarization of the observed outbursts was not measured, but observations of Payne-Scott(15) indicate that sporadic bursts (such as outbursts) are not, in general, circularly polarized. At present, therefore, this explanation of multi-band spectra in terms of magnetic fields is uncertain.

(b) *Possible Identification of the Supposed Moving Disturbances :  
Surge Prominences and Magnetic Storm Particles*

If, as suggested above, the outbursts are caused by disturbances moving outwards through the solar atmosphere, it is of interest to inquire whether the disturbances can be identified with any phenomena known to take place on the sun at the time of the flares. Two possibilities suggest themselves (see 4, 16) : the "surge" prominences that are observed on the sun at the time of flares (see, for instance, 17) ; and the particles, known to be emitted from the sun at the time of flares, that cause terrestrial magnetic storms with sudden commencements about a day later(18). The former attain radial velocities of about 100–200 km./sec. ; the latter travel between the sun and earth at mean velocities of about 800–1600 km./sec. Both are of the same order as the calculated velocities associated with outbursts (see Fig. 2).

Two pieces of evidence may be significant in any attempted identification : firstly, the fact that frequency drift has been observed only in the one direction corresponding to *outward* motion ; and secondly, the indication that the typical drift corresponds to acceleration outwards (as shown in Fig. 2). The suggestion is therefore of an accelerating outrush of matter that escapes completely from the sun. This evidence does not support the notion that the agency corresponds to *visible* matter seen in a surge prominence, the leading edge of which normally travels out to a height of about  $10^5$  km. above the photosphere and then falls back to the sun. It may, however, correspond to matter which escapes visual detection because of its high degree of ionization. The evidence seems quite consistent, both qualitatively and quantitatively, with the hypothesis that the agency corresponds to the particles that cause the terrestrial magnetic storms.

In conclusion it should be remarked that only one interpretation of the observed outburst spectra has been examined in this discussion. If this interpretation is right, or if, in general, we can associate the wave frequency of certain spectral features with the locality of disturbances in the solar atmosphere, then the study of outburst spectra seems likely to offer data of considerable value to solar research.

## VI. ACKNOWLEDGMENTS

The author is indebted to Mr. L. L. McCready, Mr. J. D. Murray, and Mr. W. C. Rowe who took part in the observational phase of this work. The author also wishes to thank Dr. J. L. Pawsey and Mr. S. F. Smerd for helpful criticism and advice in the preparation of this paper.

## VII. REFERENCES

- (1) APPLETON, E. V., and HEY, J. S.—*Phil. Mag.* **37** : 73-84 (1946).
- (2) LOVELL, A. C. B., and BANWELL, C. J.—*Nature* **158** : 517-8 (1946).
- (3) ALLEN, C. W.—*Mon. Not. R. Astr. Soc.* **107** : 386-96 (1947).
- (4) PAYNE-SCOTT, RUBY, YABSLEY, D. E., and BOLTON, J. G.—*Nature* **160** : 256-7 (1947).
- (5) RYLE, M., and VONBERG, D. D.—*Proc. Roy. Soc. A* **193** : 98-120 (1948).
- (6) COVINGTON, A. E.—*Proc. Inst. Radio Engrs. N.Y.* **36** : 454-7 (1948).
- (7) HEY, J. S., PARSONS, S. J., and PHILLIPS, J. W.—*Mon. Not. R. Astr. Soc.* **108** : 354-71 (1948).

- (8) WILD, J. P., and MCCREADY, L. L.—*Aust. J. Sci. Res. A* **3** : 387 (1950).
- (9) MARTYN, D. F.—*Nature* **159** : 26-7 (1947).
- (10) BAILEY, V. A.—*Phys. Rev.* **78** : 428-43 (1950).
- (11) ALLEN, C. W.—*Mon. Not. R. Astr. Soc.* **107** : 426-32 (1947).
- (12) JAEGER, J. C., and WESTFOLD, K. C.—*Aust. J. Sci. Res. A* **2** : 322-34 (1949).
- (13) WESTFOLD, K. C.—*Aust. J. Sci. Res. A* **2** : 169-83 (1949).
- (14) CHAPMAN, S.—*Mon. Not. R. Astr. Soc.* **103** : 117 (1943).
- (15) PAYNE-SCOTT, RUBY.—*Aust. J. Sci. Res. A* **2** : 214-27 (1949).
- (16) PAWSEY, J. L.—Solar radio-frequency radiation. *Proc. Inst. Elect. Engrs.* **97** (1950).
- (17) ELLISON, M. A.—*Mon. Not. R. Astr. Soc.* **109** : 3-27 (1949).
- (18) NEWTON, H. W.—*Mon. Not. R. Astr. Soc.* **103** : 244-57 (1943).

### EXPLANATION OF PLATES 1, 2

The plates show the dynamic spectrum of the starts of four outbursts in the frequency range 70–130 Mc/s. Intensities are shown as contours in the time-frequency plane. The intensity contour scale is logarithmic and is shown in the accompanying key.

#### PLATE 1

- (a) Outburst No. 1. Time of start : July 27, 1949, 02 hr. 04 min. (G.M.T.).
- (b) Outburst No. 2. Time of start : April 6, 1949, 05 hr. 04 min.

#### PLATE 2

- (a) Outburst No. 3. Time of start : June 8, 1949, 02 hr. 17 min.
- (b) Outburst No. 4. Time of start : February 14, 1949, 01 hr. 32 min.

(b) *Possible Identification of the Supposed Moving Disturbances :  
Surge Prominences and Magnetic Storm Particles*

If, as suggested above, the outbursts are caused by disturbances moving outwards through the solar atmosphere, it is of interest to inquire whether the disturbances can be identified with any phenomena known to take place on the sun at the time of the flares. Two possibilities suggest themselves (see 4, 16) : the "surge" prominences that are observed on the sun at the time of flares (see, for instance, 17) ; and the particles, known to be emitted from the sun at the time of flares, that cause terrestrial magnetic storms with sudden commencements about a day later (18). The former attain radial velocities of about 100–200 km./sec. ; the latter travel between the sun and earth at mean velocities of about 800–1600 km./sec. Both are of the same order as the calculated velocities associated with outbursts (see Fig. 2).

Two pieces of evidence may be significant in any attempted identification : firstly, the fact that frequency drift has been observed only in the one direction corresponding to *outward* motion ; and secondly, the indication that the typical drift corresponds to acceleration outwards (as shown in Fig. 2). The suggestion is therefore of an accelerating outrush of matter that escapes completely from the sun. This evidence does not support the notion that the agency corresponds to *visible* matter seen in a surge prominence, the leading edge of which normally travels out to a height of about  $10^5$  km. above the photosphere and then falls back to the sun. It may, however, correspond to matter which escapes visual detection because of its high degree of ionization. The evidence seems quite consistent, both qualitatively and quantitatively, with the hypothesis that the agency corresponds to the particles that cause the terrestrial magnetic storms.

In conclusion it should be remarked that only one interpretation of the observed outburst spectra has been examined in this discussion. If this interpretation is right, or if, in general, we can associate the wave frequency of certain spectral features with the locality of disturbances in the solar atmosphere, then the study of outburst spectra seems likely to offer data of considerable value to solar research.

## VI. ACKNOWLEDGMENTS

The author is indebted to Mr. L. L. McCready, Mr. J. D. Murray, and Mr. W. C. Rowe who took part in the observational phase of this work. The author also wishes to thank Dr. J. L. Pawsey and Mr. S. F. Smerd for helpful criticism and advice in the preparation of this paper.

## VII. REFERENCES

- (1) APPLETON, E. V., and HEY, J. S.—*Phil. Mag.* **37** : 73-84 (1946).
- (2) LOVELL, A. C. B., and BANWELL, C. J.—*Nature* **158** : 517-8 (1946).
- (3) ALLEN, C. W.—*Mon. Not. R. Astr. Soc.* **107** : 386-96 (1947).
- (4) PAYNE-SCOTT, RUBY, YABSLEY, D. E., and BOLTON, J. G.—*Nature* **160** : 256-7 (1947).
- (5) RYLE, M., and VONBERG, D. D.—*Proc. Roy. Soc. A* **193** : 98-120 (1948).
- (6) COVINGTON, A. E.—*Proc. Inst. Radio Engrs. N.Y.* **36** : 454-7 (1948).
- (7) HEY, J. S., PARSONS, S. J., and PHILLIPS, J. W.—*Mon. Not. R. Astr. Soc.* **108** : 354-71 (1948).



- (8) WILD, J. P., and MCCREADY, L. L.—*Aust. J. Sci. Res. A* **3**: 387 (1950).
- (9) MARTYN, D. F.—*Nature* **159**: 26-7 (1947).
- (10) BAILEY, V. A.—*Phys. Rev.* **78**: 428-43 (1950).
- (11) ALLEN, C. W.—*Mon. Not. R. Astr. Soc.* **107**: 426-32 (1947).
- (12) JAEGER, J. C., and WESTFOLD, K. C.—*Aust. J. Sci. Res. A* **2**: 322-34 (1949).
- (13) WESTFOLD, K. C.—*Aust. J. Sci. Res. A* **2**: 169-83 (1949).
- (14) CHAPMAN, S.—*Mon. Not. R. Astr. Soc.* **103**: 117 (1943).
- (15) PAYNE-SCOTT, RUBY.—*Aust. J. Sci. Res. A* **2**: 214-27 (1949).
- (16) PAWSEY, J. L.—Solar radio-frequency radiation. *Proc. Inst. Elect. Engrs.* **97** (1950).
- (17) ELLISON, M. A.—*Mon. Not. R. Astr. Soc.* **109**: 3-27 (1949).
- (18) NEWTON, H. W.—*Mon. Not. R. Astr. Soc.* **103**: 244-57 (1943).

### EXPLANATION OF PLATES 1, 2

The plates show the dynamic spectrum of the starts of four outbursts in the frequency range 70–130 Mc/s. Intensities are shown as contours in the time-frequency plane. The intensity contour scale is logarithmic and is shown in the accompanying key.

#### PLATE 1

- (a) Outburst No. 1. Time of start: July 27, 1949, 02 hr. 04 min. (G.M.T.).
- (b) Outburst No. 2. Time of start: April 6, 1949, 05 hr. 04 min.

#### PLATE 2

- (a) Outburst No. 3. Time of start: June 8, 1949, 02 hr. 17 min.
- (b) Outburst No. 4. Time of start: February 14, 1949, 01 hr. 32 min.

- (8) WILD, J. P., and McCREADY, L. L.—*Aust. J. Sci. Res. A* **3**: 387 (1950).
- (9) MARTYN, D. F.—*Nature* **159**: 26-7 (1947).
- (10) BAILEY, V. A.—*Phys. Rev.* **78**: 428-43 (1950).
- (11) ALLEN, C. W.—*Mon. Not. R. Astr. Soc.* **107**: 426-32 (1947).
- (12) JAEGER, J. C., and WESTFOLD, K. C.—*Aust. J. Sci. Res. A* **2**: 322-34 (1949).
- (13) WESTFOLD, K. C.—*Aust. J. Sci. Res. A* **2**: 169-83 (1949).
- (14) CHAPMAN, S.—*Mon. Not. R. Astr. Soc.* **103**: 117 (1943).
- (15) PAYNE-SCOTT, RUBY.—*Aust. J. Sci. Res. A* **2**: 214-27 (1949).
- (16) FAWSEY, J. L.—Solar radio-frequency radiation. *Proc. Inst. Elect. Engrs.*
- (17) ELLISON, M. A.—*Mon. Not. R. Astr. Soc.* **109**: 3-27 (1949).
- (18) NEWTON, H. W.—*Mon. Not. R. Astr. Soc.* **103**: 244-57 (1943).

#### EXPLANATION OF PLATES 1, 2

The plates show the dynamic spectrum of the starts of four outbursts in the range 70-130 Mc/s. Intensities are shown as contours in the time-frequency plane. The scale is logarithmic and is shown in the accompanying key.

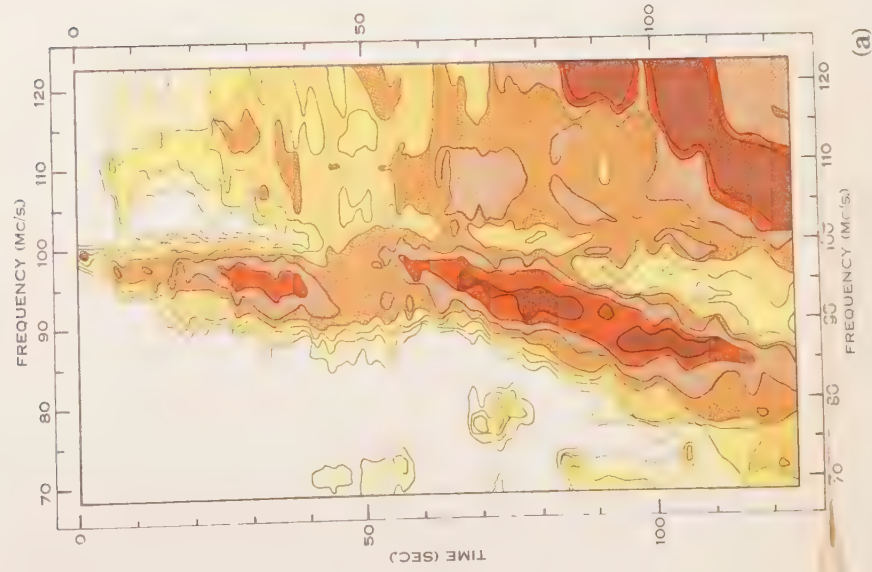
##### PLATE 1

- a Outburst No. 1. Time of start: July 27, 1949, 02 hr. 04 min. (G.M.T.).
- b Outburst No. 2. Time of start: April 6, 1949, 05 hr. 04 min.

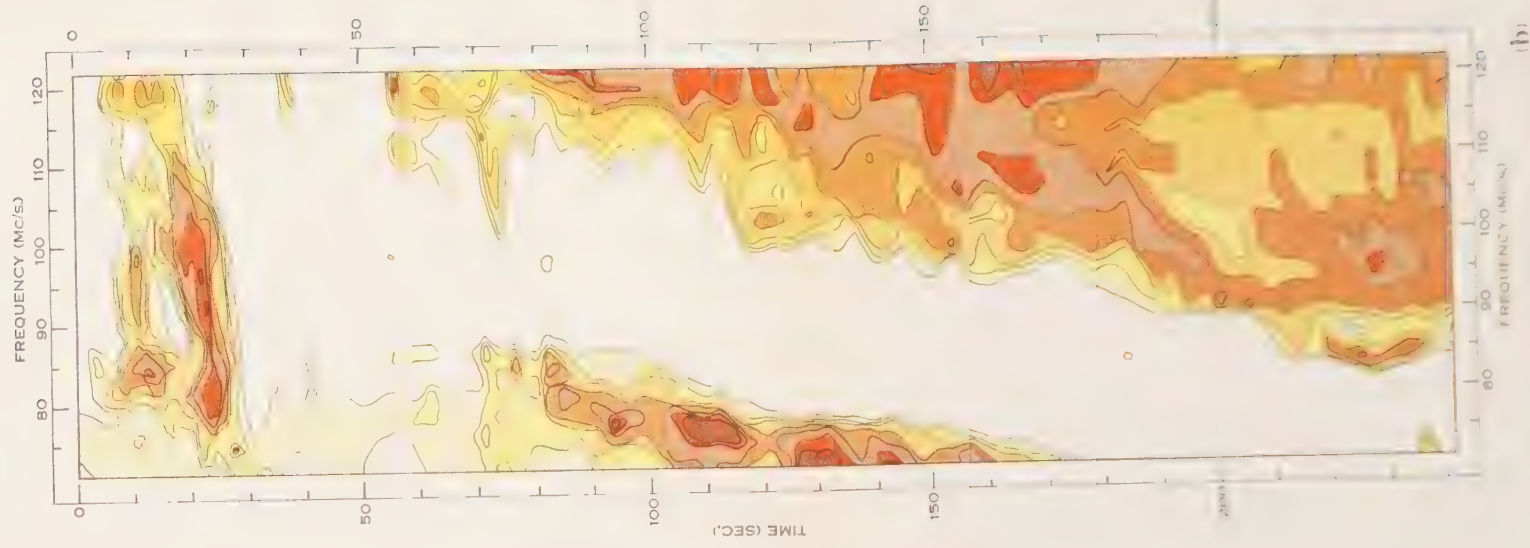
##### PLATE 2

- a Outburst No. 3. Time of start: June 8, 1949, 02 hr. 17 min.
- b Outburst No. 4. Time of start: February 14, 1949, 01 hr. 32 min.





(a)



(b)

AUST. J. SCI. RES., A, VOL. 3, 1950

PLATE 2

WILD.—OBSERVATIONS OF THE SPECTRUM OF HIGH-INTENSITY  
SOLAR RADIATION AT METRE WAVELENGTHS. II





# ATMOSPHERIC NOISE LEVELS AT RADIO FREQUENCIES NEAR DARWIN, AUSTRALIA

By D. E. YABSLEY\*

[*Manuscript received February 27, 1950*]

## *Summary*

Between August 25, 1944, and October 31, 1945, a practically continuous record of the average level of atmospheric radio noise at a frequency of 1.93 Mc/s. was obtained near Darwin, in north-western Australia. A few measurements were also made at a frequency of 5.9 Mc/s.

The noise-measuring programme is described and the results obtained are presented graphically.

## I. INTRODUCTION

In August 1944, the taking of measurements of atmospheric noise levels at radio frequencies was begun near Darwin by the Operational Research Group of the U.S. Signal Corps. As the result of an agreement with the Division of Radiophysics, C.S.I.R.O., and the Royal Australian Air Force, the measurements were continued for a year, the Royal Australian Air Force providing the necessary staff for the operation and maintenance of the equipment from October 1944 onwards. A practically continuous record of the average level of atmospheric noise on a frequency of 1.93 Mc/s. between August 1944 and October 1945 was obtained, together with a limited number of measurements at 5.9 Mc/s.

The results of this project have been given previously in restricted laboratory reports(1-3) which summarized the measurements made in 1944 and tabulated the hourly mean noise levels up to the end of August 1945. This paper reports in a condensed form all the results obtained.

## II. EQUIPMENT

The noise-measuring equipment consisted of the following items:

(1) A vertical aerial, about 40 ft. high, and a ground system consisting of 11 wires each 250 ft. long, laid along the surface of the earth, radially from the hut in which the equipment was housed.

(2) Two communications-type receivers, modified for recording, and two recording milliammeters. The receiver modifications consisted of the addition of a cathode-follower stage, driven by the output of the diode second detector. The recording milliammeter formed part of the cathode-follower load. The grid circuit time-constant was 40 seconds, so that a continuous smoothing process was carried out over this interval, and the average level was indicated on the recorder. Facilities were provided for aural monitoring of the receiver output,

\* Division of Radiophysics, C.S.I.R.O., University Grounds, Sydney.

and for shortening the time-constant of the recording stage for calibration purposes. The receiver bandwidths which were calibrated (see below) were about 5 kc/s.

(3) A signal generator and a dummy aerial, with the aid of which the receiver response was calibrated daily.

### III. THE SPECIFICATION AND INTERPRETATION OF MEASURED NOISE LEVELS

In order to calibrate the system a standard field was produced at the receiving site by setting up a transmitter at a convenient distance. This field was measured with an RCA type 308A Field Strength Meter. This standard signal was also received on the noise-recording receiver, and the response noted. The input was then switched to the dummy aerial and signal generator, and the output of the latter adjusted to give the same response. This gave one point on the curve relating field strength and receiver output, and the signal generator was used to obtain the remainder of the curve.

The signal generator was also used as a reference standard for day-to-day checks of receiver calibration.

The directional diagram of the aerial in the horizontal plane was checked by repeating the calibration test with the transmitter at different azimuths. The maximum variations from the mean value observed were about one decibel.

The power bandwidth of the receiver was determined by integrating the curve of the receiver response against frequency. The values so obtained were 5.7 and 4.2 kc/s. for the 1.93 and 5.9 Mc/s. receivers respectively. The bandwidths of both receivers between the -6 db. points were 8 kc/s.

If the received noise voltage were truly random and if the received intensity were  $f$  microvolts per metre, then the field strength of the noise contained in a bandwidth of 1 kc/s. would be  $f/\sqrt{B_{kc}}$  where  $B_{kc}$  is the power bandwidth in kilocycles per second.

It may be objected that the noise voltage will not be truly random except when the only thunderstorms operative are distant, and that the closer the storms are situated the more impulsive will be the received noise voltage. The measured value of noise field strength under these conditions would tend towards a direct proportionality to the receiver bandwidth, instead of being proportional to the square root of it. However, in the receiving system used for the present series of observations, there was the further limitation that high noise peaks, such as those due to lightning flashes from local storms, caused overloading of the receiver. This was inevitable because it was not possible to record a range of field strengths exceeding about 45 db. without undue compression of the low end of the scale. Thus the recording of the higher noise intensities was subject to two major difficulties, but it may be noted that if the apparent intensities are reduced to equivalent values for 1 kc. bandwidth by dividing by  $\sqrt{B_{kc}}$  the errors in the highest intensities tend to cancel one another. For this reason it was thought justifiable to rationalize the unit of noise field strength in this manner.

Another important practical problem which is related to the above considerations is that of estimating from measured noise levels the minimum field strengths required for communication purposes. This is difficult because the required signal-to-noise ratio for any particular service does not depend solely on one variable such as the average noise field strength; approximate values may, however, be calculated from the following signal-to-noise ratio quoted in reference (1):

For 90 per cent. intelligibility of related word messages a carrier field strength of 12 times the average field strength of atmospheric noise in a 1 kc. band is required. (The peak modulation is assumed to be 100 per cent.)

Factors needed to modify the values of required field intensity given by the above formula for other types of service are given in a circular of the United States National Bureau of Standards(4).

#### IV. THE NOISE-MEASURING PROGRAMME

Measurements on a frequency of 1.93 Mc/s. were made by the U.S. Signal Corps at a flat site on the coast a few miles from Darwin from August 25 to September 25, 1944. They were continued by the Division of Radiophysics, C.S.I.R.O., and the Royal Australian Air Force at a site in moderately timbered, undulating country about 50 miles inland from Darwin. (Most of the trees in the vicinity of the site were about 30 ft. high.) Recording began here on October 21, 1944 and was practically continuous until October 31, 1945. No interference from signals was encountered at this frequency.

Interference was troublesome at a frequency of 5.9 Mc/s., however, and blanketed the noise on most days on which recording was attempted. Long-term continuous recording was not tried at this frequency, records being obtained on only three days in November and December 1944 and on most days between May 15 and July 31, 1945.

#### V. RESULTS

##### (a) *Results at 1.93 Mc/s.*

The results at this frequency are shown in Figures 1 to 4.

The median values of noise field strength for each hour of each day were first calculated, and then monthly-mean values of field strength for each hour of the day were computed. These are shown for the whole period of observation in the single contour diagram (Fig. 1). It should be noted that the minimum field strength which could be measured accurately was 0.1 microvolt per metre per kilocycle bandwidth. Between March and September the noise level in the middle of the day almost always fell below this field strength to an unknown value.

All significant seasonal variations in noise level in this area may be described by dividing the year into two main seasons, with brief transition periods between :

(1) From the beginning of November to the end of February.—This is the period of frequent local thunderstorms associated with the wet north-west monsoon.



(2) From March to September.—This period includes all of the dry season of south-east winds and fine clear weather, during which the nearest thunderstorm belts are at least 1000 miles away to the north-east and north-west. The transition from noise conditions characteristic of the wet season to those of the dry season appeared to be very sharp, and in 1945 occurred around the beginning of March.

(3) The month of October.—In both 1944 and 1945 this was a month of transition between the two main seasons, during which afternoon thunderstorms became common.

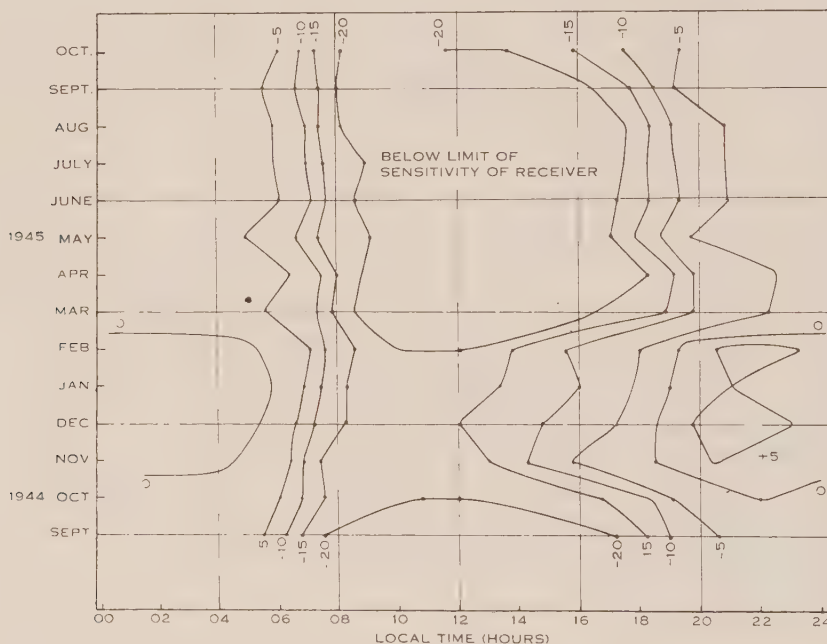


Fig. 1.—Contours of the average field strength of atmospheric radio noise at 1.93 Mc/s. for each month of the year and each hour of the day. The units are decibels above one microvolt per metre for one kilocycle bandwidth. The limit of sensitivity of the receiver was about  $0.1 \mu\text{V.}/\text{m.}/(\text{kc/s.})^{\frac{1}{2}}$ , i.e.  $-20 \text{ db}$ .

Figure 2 shows the three curves obtained by taking the means, over the above groups of months, of the monthly-mean level of noise at each hour of the day. The main features shown by these curves are clearly correlated with the fact that an important thunderstorm belt is centred on the Darwin area during the southern summer—the wet season—while there is none within a range of 1000 miles of Darwin during the northern summer—the dry season.

In the dry season the day-time noise level was low, owing to poor propagation conditions and the absence of local thunderstorms. The noise level was higher at night owing to more favourable sky-wave propagation, and was fairly steady. The steadiness was probably due to the great distance from the receiver of the major noise sources, resulting in roughly equal contributions from a large number of thunderstorms.

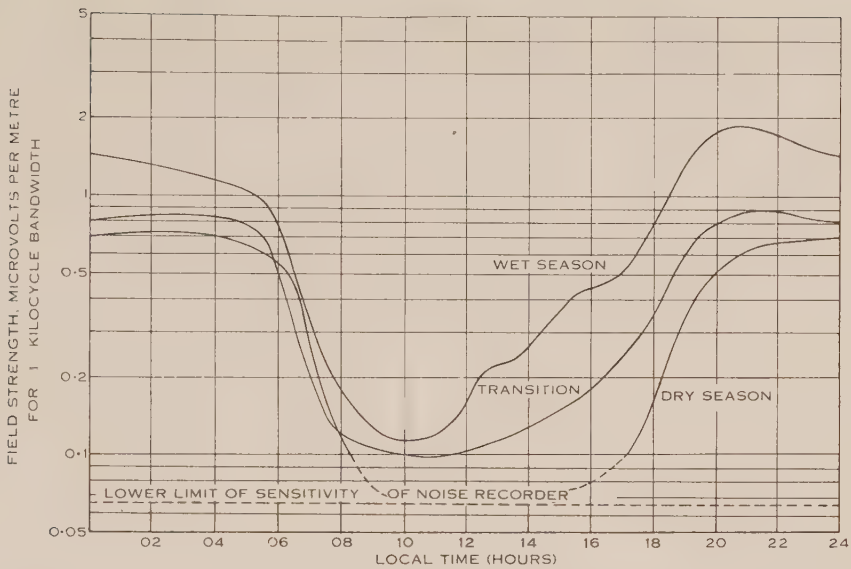


Fig. 2.—Curves showing means of monthly-mean field strengths of atmospheric radio noise at 1.93 Mc/s. at each hour of the day.

*Wet season*: November 1, 1944, to February 28, 1945.

*Dry season*: August 25 to September 25, 1944, and March 1 to September 30, 1945.

*Transition period*: October 21–31, 1944, and October 1–31, 1945.

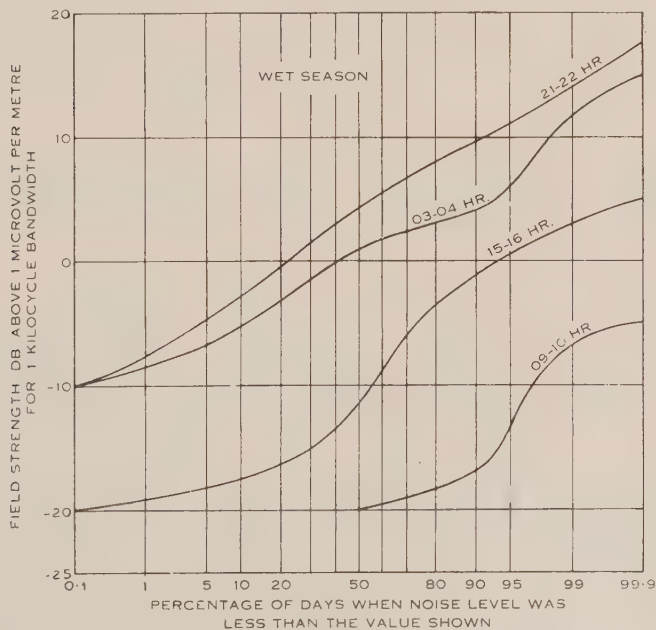


Fig. 3.—Curves showing the distribution of the observations of field strength of atmospheric noise at 1.93 Mc/s., at selected hours of the day during the period from November 1, 1944, to February 28, 1945.

In the wet season there was a minimum noise level in the forenoon, but local storms caused a moderate rise in level during the afternoon, and also resulted in the peak intensity being recorded in the early evening. The seasonal variation in the night-time noise level was around 6 db.

Figures 1 and 2 show mean noise levels. It is equally desirable to know the variability of the noise field strength. This factor may be determined from Figures 3 and 4, which show the distribution of the individual observations at selected hours of the day and night during the two main noise seasons. The

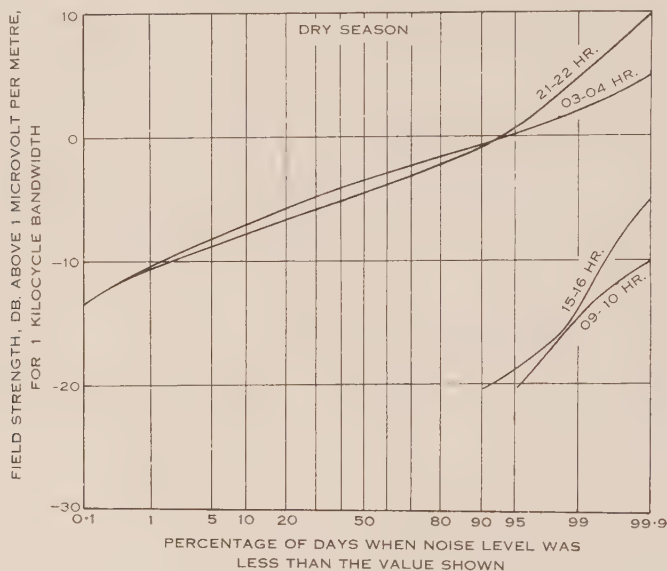


Fig. 4.—Curves showing the distribution of the observations of field strength of atmospheric noise at 1.93 Mc/s. at selected hours of the day during the period from August 25 to September 25, 1944, and March 1 to September 30, 1945.

curves give the percentage of the days on which the noise level was below any chosen level. The abscissa scale is a "normal probability scale" which is ruled in such a way that a series of intensities whose logarithms showed a normal distribution, would be represented by a straight line. The individual night-time noise field strengths showed an approximately normal distribution during both seasons. The occurrence of local storms only in the wet season resulted in a seasonal change in variability. In the dry season 50 per cent. of the observations were included in a range of  $\pm 1.7$  db. from the median, while the corresponding range was  $\pm 3$  db. in the wet season.

During the wet season the day-time observations showed marked departures from a normal distribution owing to the intermittent occurrence of local thunderstorms. The finite sensitivity of the receiver limits the significance of the day-time dry season observations.

*(b) Results at 5.9 Mc/s.*

The curve obtained by taking the mean noise level at each hour over the three days on which records were obtained during the wet season is shown in Figure 5. Comparison of the observed noise levels at 1.93 Mc/s. on the same three days with the mean levels for November and December shows that these days were ones of slightly less than average noise activity. It is probable that this result also holds at 5.9 Mc/s.

During the period from May 15 to July 31, 1945, continuous recording was attempted, but interference from signals was very troublesome, the record being lost for over 50 per cent. of the time, between 23 hours and 09 hours local time.

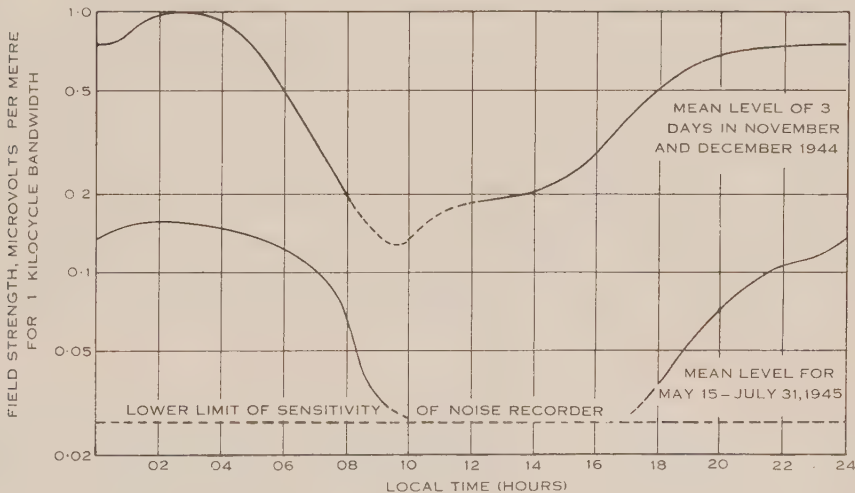


Fig. 5.—Curves showing mean recorded field strengths of atmospheric radio noise at 5.9 Mc/s.

The times at which interference was bad probably corresponded to good propagation conditions, and therefore probably also to higher than average noise levels. Accordingly, the average recorded levels may be lower than the true average for these two and a half months. There was no significant difference between the average levels recorded in the separate months. Figure 5 shows the mean levels at each hour of the day over the whole period. The day-time level during this period was below the minimum intensity which could be recorded. The seasonal variation of night-time noise level on this frequency was about 15 db., against 6 db. at 1.93 Mc/s.

*(c) Comparison between Observed and Expected Noise Levels*

The National Bureau of Standards circular already referred to(4) gives predictions of minimum field intensities required to assure communication 90 per cent. of the time in the presence of atmospheric noise. By allowing for the specified bandwidth and signal-to-noise ratio the assumed noise levels can be determined. The noise intensities so obtained for the Darwin area at 1.93 Mc/s. were compared with the corresponding measured intensities obtained from



Figures 3 and 4. This showed that while the assumed night-time noise levels were too high in all seasons by factors of from 8 to 15 db., they were about 6 db. low in the daytime of the wet season.

The observations at 5.9 Mc/s. were limited in number and the sampling was poor. Consequently a comparison with the expected noise levels was not considered worth while.

## VI. ACKNOWLEDGMENTS

This work formed part of the wartime programme of research carried out by the Division of Radiophysics, C.S.I.R.O. It was planned and executed in collaboration with the Interservice Radio Propagation Laboratory, Washington, the United States Signal Corps, and the Royal Australian Air Force. Thanks are due to the United States Signal Corps for the major part of the equipment, to the Postmaster-General's Department for the loan of a recording milliammeter, and particularly to the Royal Australian Air Force for the provision of the necessary personnel for the erection and operation of the noise-measuring station.

Thanks are also due to Mr. R. Bateman for making available the results of the initial observations made by him during August and September 1944 for the Operational Research Branch of the United States Signal Corps.

The author also wishes to thank Dr. J. L. Pawsey for many helpful discussions at various stages of the project.

## VII. REFERENCES

- (1) BATEMAN, R., HERBSTREIT, J. W., and ZECHIEL, R. B.—Measurement of factors affecting radio communication and Loran navigation in the south-west Pacific area. Operational Research Branch, Office of Chief Signal Officer, Washington, Rep. ORB-2-4 (1944).
- (2) YABSLEY, D. E.—Atmospheric noise measurements in Darwin. Div. Radiophysics, Coun. Sci. Industr. Res. Aust. Rep. RP 235/1 (1945).
- (3) YABSLEY, D. E.—Supplementary bulletins to "Atmospheric noise measurements in Darwin". Div. Radiophysics, Coun. Sci. Industr. Res. Aust. Rep. RP 235/2-9 (1945).
- (4) Ionospheric radio propagation. U.S. Dep. Comm. Nat. Bur. Stand. Circ. 462 (1948).

# THE PRESSURE DEPENDENCE OF THE THERMAL CONDUCTIVITY OF POLYATOMIC GASES AT 0 °C.

By W. G. KANNULUIK\* and HEATHER B. DONALD\*

[*Manuscript received May 29, 1950*]

## *Summary*

A form of "hot wire" apparatus is described in which the hot wire is replaced by a thin-walled nickel tube  $\frac{1}{8}$  in. in diameter mounted in a similar nickel tube  $\frac{1}{4}$  in. in diameter. The gas under investigation occupies the narrow annular space between the two tubes.

It is shown that the apparatus is completely free from convection up to pressures somewhat greater than 1 atm. and that the temperature discontinuity effect is inappreciable down to pressures of a few cm. of mercury. The apparatus is used to test the pressure dependence of the thermal conductivity of several polyatomic gases.

The thermal conductivity (at 0 °C.) of the gases carbon dioxide and methane is found to increase slowly with the pressure over the range 10–80 cm. of mercury. It is suggested that this effect may be taken as evidence of the incomplete participation of the energy of vibration of the molecules in the conduction process.

## I. INTRODUCTION

The present paper describes experiments primarily carried out to investigate the dependence of thermal conductivity of polyatomic gases on pressure. Owing to the incomplete exchange of energy between the translational and vibrational energy of the molecules of a polyatomic gas (carbon dioxide) Ubbelohde(1) has shown on theoretical grounds that the conductivity should decrease with the pressure but at sufficiently high pressures when the exchange of energy becomes complete it should approach a constant value independent of the pressure. Ubbelohde studied the case of a gas enclosed between parallel plates and deduced that the pressure dependence would be quite small at ordinary pressures.

While the parallel plate method has the advantages of being completely free from convection and from the effect of temperature discontinuity at the plate surfaces, this method is less sensitive to small changes in the conductivity than is the "hot wire" method. It was decided, therefore, to modify the "hot wire" method used in this Laboratory over a number of years so that it would conform more nearly to the conditions of the parallel plate method. The earlier types of apparatus(2) consisting of a short thick wire mounted coaxially in a metal tube have been replaced in the present work by a short thin-walled nickel tube mounted coaxially in a somewhat wider similar nickel tube, the gas being contained in the narrow annular gap (0.15 cm. wide) between the two tubes. This modification leaves the existing theory developed for the short thick wire apparatus intact but has the advantages of the complete absence of

\* Physics Department, University of Melbourne.

convection at pressures at least as high as 1 atm. and of insensitivity to the temperature discontinuity effect until the pressure is reduced to a few cm. of mercury.

It should be pointed out that the older traditional forms of the "hot wire" experiment, in which a relatively long fine filament is employed, cannot be used to obtain direct evidence of the pressure dependence of the conductivity since in this type of apparatus the effect of the temperature discontinuity is appreciable even at ordinary pressures. Nevertheless the attempt has been made recently by Schäfer, Rating, and Eucken(3) using an apparatus of this type to separate the Ubbelohde effect from the temperature discontinuity effect. By comparing the experimental results obtained for a monatomic gas (argon) with those obtained for a polyatomic gas (carbon dioxide), they tried to separate the effect of the suppressed energy exchange between the translational and the vibrational energy from the normal pressure fall due to the temperature discontinuity. In this way it was hoped to estimate the degree to which the vibrational energy is involved in the conduction process. The disadvantage of this procedure is that the two effects are interwoven and there is no independent means of determining either of the two accommodation coefficients involved. The best that can be done is to select values of the accommodation coefficients which best fit the experimental data.

## II. DESCRIPTION OF THE "HOT WIRE" APPARATUS

The apparatus was constructed out of thin-walled cold drawn nickel tubing. Two short lengths of tubing  $\frac{1}{8}$  in. and  $\frac{1}{4}$  in. in external diameter were selected for straightness and uniformity. The piece of  $\frac{1}{8}$  in. tubing replaced the thick wire used in previous types of apparatus. This was annealed at a temperature of 400 °C. before assembly.

The construction of the apparatus is shown in Figure 1. The inner tube  $T$  is mounted coaxially in the outer tube  $T'$  and is electrically insulated from the latter by means of a copper-glass-copper seal  $S$ . The inner tube passes through and is soft soldered to a copper cap at the lower end and a copper plug  $G$  at the upper end. The plug  $G$  has three spaced vertically bored holes as shown in plan. The narrow annular space between the two tubes contains the gas under investigation and may be put into communication with the pumping system or the gas supply.

The inner tube  $T$  is sealed at both ends, thus leaving it filled with stagnant air. It was assumed that the presence or absence of gas in the inner tube would not affect the longitudinal distribution of temperature along it.

Heating currents of the order of a few amperes maintained by a small battery of storage cells were passed through the inner tube via the current leads  $C$  and  $C'$ . A pair of potential leads  $p$  and  $p'$  soldered to the inner tube were so placed that the distance between them could be accurately measured.

### (a) *Dimensions of the Apparatus*

The metrology of the apparatus was carried out with a micrometer screw gauge and a Mercer dial gauge. The screw gauge was calibrated against a set of Tower slip gauges.

The principal requirements are :

- (1) The inner  $\frac{1}{8}$  in. tube should be straight, and of constant external diameter and wall thickness.
- (2) The inner diameter of the  $\frac{1}{4}$  in. outer tube should be constant along its length.

Tests were made on short lengths of tube (2 in. long) cut off from the ends of the two selected tubes. The test pieces were cut lengthwise into halves and

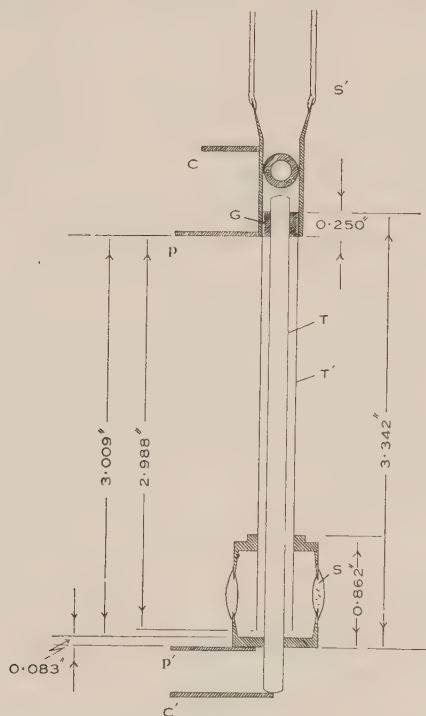


Fig. 1.—The conductivity apparatus.

were tested for constancy and uniformity of wall thickness by means of a screw gauge with small half spheres wrung on to the jaws. The thickness of the test pieces was found to be constant and uniform and accordingly it was assumed that the wall thickness of the two tubes was also constant. The inner diameter of the outer tube was obtained by subtracting twice the wall thickness from the measured external diameter. It is not necessary to know the inner diameter of the inner tube.

The following are the principal dimensions of the apparatus :

Mean external diameter of the outer tube :	0.2525 in.
Mean wall thickness :	0.0051 in.
Mean internal diameter of outer tube :	2423 in. = $0.6154 \pm 0.0002$ cm.
Mean external diameter of the inner tube :	$0.1251 \text{ in.} = 0.3178 \pm 0.0002 \text{ cm.}$
Length (2 <i>l</i> ) of inner tube (Fig. 1) :	$3.009 \text{ in.} = 7.643 \pm 0.001 \text{ cm.}$



### III. ELIMINATION OF CONVECTION FROM "HOT WIRE" APPARATUS

In any "hot wire" apparatus the complete absence of convection over a range of gas pressures is essential. A systematic study of convection in a gas enclosed between concentric cylinders has been made by Weber(4). He proved that convection was minimized by placing the axis of the apparatus vertical and by making the annular gap containing the gas as narrow as possible. If these conditions were complied with there would be considerable resistance to any mass movement of the gas and convection would be confined to the ends of the apparatus.

A quantitative study of convection in a gas contained between concentric cylinders has been made by Ulsamer(5). To take into account the effect of convection Ulsamer introduced a quantity  $\zeta$  defined as the ratio of the apparent conductivity  $k_{app}$  of the gas to its real conductivity  $k$ . It follows that

$$\zeta = \frac{k_{app}}{k} \geq 1.$$

Ulsamer also expressed  $\zeta$  as a function of the product of two quantities, that is,

$$\zeta = f(Gr.Pr),$$

where  $Gr = \delta^3 g \beta (t_i - t_a) / \nu^2$ ,

$$Pr = \nu / a,$$

and  $\delta$  = annular gap,  $g$  the acceleration due to gravity,  $\beta$  the coefficient of expansion of the gas,  $(t_i - t_a)$  the temperature drop across the annulus,  $\nu$  the kinematic viscosity or  $\eta / \rho$ , where  $\eta$  is the viscosity and  $\rho$  the density of the gas,

and  $a$  is the thermometric conductivity, or  $\frac{k}{\rho c_v}$  the thermal conductivity of the gas divided by the product of the density and the isometric specific heat.

From a careful examination of the available data Ulsamer was enabled to formulate a criterion for a vertical hot wire apparatus to be free from convection to within 1 per cent., that is, that  $k_{app}/k \geq 1.01$ . The criterion is that the product  $Gr.Pr$  should not exceed the numerical value of 2000.

The quantity  $Gr$  is seen to be proportional to the cube of the annular gap  $\delta$  and directly proportional to the temperature drop over the gap. Accordingly in designing a "hot wire" apparatus the annular gap should be made as small as possible. In the apparatus employed in the present work the annular gap was only 0.15 cm. and the average temperature drop of the order of 2 °C. For air at atmospheric pressure the product  $Gr.Pr$  was approximately equal to unity.\*

It follows from the definition of the quantities entering into  $Pr$ , that  $Pr$  is also proportional to  $\rho^2$  or to the square of the density of the gas. On the other hand  $Pr$  is independent of the density. Hence the product  $Gr.Pr$  is proportional to the square of the density. This implies that if convection is present in an apparatus its effect will be diminished by reducing the pressure of the gas, that is  $k_{app}$  will decrease with decreasing gas pressure. If  $k_{app}$  is constant over a range of pressure there can be no convection present over this range of

\* As Ulsamer's criterion ( $Gr.Pr \geq 2000$ ) is so much exceeded in our apparatus ( $Gr.Pr = 1$ ) it can be expected that the present apparatus will be entirely free from convection.

pressure. This appears to be the only safe criterion for the complete absence of convection in a "hot wire" apparatus, and was complied with strictly in the apparatus used in the present work.

#### IV. THEORY

The complete theory of the "hot wire" method used in the present work has been previously given in detail(6, 7) and will not be repeated here. If a wire of thermal conductivity  $\lambda$  is mounted along the axis of a tube maintained everywhere at  $0^\circ\text{C}$ . and the tube is filled with a gas of thermal conductivity  $k$ , then we have

$$\lambda\pi b^2 \frac{\partial^2 \theta}{\partial z^2} + 2\pi b k \left( \frac{\partial \theta}{\partial r} \right)_{r=b} + \frac{I^2 \rho_0}{J} (1 + \alpha \theta) = 0, \quad \dots\dots\dots (1)$$

where  $\rho = \rho_0(1 + \alpha\theta)$  is the resistance per unit length of the wire at  $\theta^\circ\text{C}$ ., and  $b$  is the radius of the wire.

If the flow of heat from the wire through the gas is strictly radial, the solution of the above equation is

$$\left. \begin{aligned} \left( \frac{1}{\beta l} \right)^2 \left( 1 - \frac{\tanh \beta l}{\beta l} \right) &= \frac{2\pi b^2 \lambda J (R - R_0)}{R_0^2 I^2 \alpha l}, \\ \beta^2 &= \frac{2h}{b\lambda} - \frac{I^2 R_0 \alpha}{2\pi b^2 J \lambda l}, \end{aligned} \right\} \dots\dots\dots (2)$$

where

and

$$h = \frac{k}{b \ln \frac{a}{b}}, \quad \dots\dots\dots (3)$$

where  $R$  is the resistance (ohm) of the wire when carrying a current  $I$  ampere,

$R_0$  is the true resistance of the wire at  $0^\circ\text{C}$ .,

$\alpha$  is the temperature coefficient, or  $\left( \frac{1}{R} \frac{dR}{d\theta} \right)_{0^\circ\text{C.}}$ ,

$J$  is the mechanical equivalent of heat,

$h$  is the external conductivity,

$2l$  is the length and  $2b$  the diameter of the wire,

$2a$  is the internal diameter of the outer tube.

When the "hot wire" apparatus is highly evacuated there remains only a very small loss of heat from the wire by radiation. We may replace the external conductivity  $h$  by  $h_R$ , where  $h_R < h$ . Relations (2) can be shown to reduce to

$$\lambda = \frac{1}{6} \frac{R R_0 I^2 \alpha l}{\pi b^2 J (R - R_0)} \left( 1 + \frac{1}{30} \frac{R_0 I^2 \alpha l}{\pi b^2 J \lambda} \right) \left( 1 - \frac{4}{5} \frac{h_R l^2}{b \lambda} \right). \quad \dots\dots (4)$$

(The simple formula  $\lambda = \frac{1}{6} \frac{R R_0 I^2 \alpha l}{\pi b^2 J (R - R_0)}$  is a close approximation to equation (4).)

By means of (4)  $\lambda$  can be determined, and on substituting the value so obtained in relations (2) these may be solved for  $h$  and the thermal conductivity  $k$  finally obtained from (3).

The replacing of the solid wire by a hollow tube does not invalidate the preceding results for the solid wire. If we treat the hollow tube as a solid wire of the same diameter as the external diameter of the tube, then the quantity  $\lambda$  obtained from (4) will not be the true conductivity of the wire. Its connection with the true conductivity  $\Omega$  is given by

$$\lambda = \frac{\Omega \pi b^2}{\pi(b^2 - c^2)}, \quad \dots \dots \dots (5)$$

where  $b$  is the external and  $c$  the internal diameter of the tube. We may proceed as if the hollow tube were a solid wire of diameter  $2b$ . The value of  $\lambda$  (now not the true conductivity) given by (4) is substituted in relations (2) and  $k$  determined just as for a solid wire. It is of course unnecessary to know the true conductivity  $\Omega$  of the tube.

The value of  $h$  deduced from (2) is subject to two small corrections: (*a*) for radiation, and (*b*) for departure from strict radial flow conditions. The value of  $h$  is reduced by approximately 1 per cent. by the radiation correction and by a little less than 1 per cent. on applying the radial flow correction. Details of these corrections were given in earlier papers(6).

## V. EXPERIMENTAL PROCEDURE

The measurements of the electrical resistance of the inner tube were made with the help of a Wolff low resistance potentiometer by direct comparison of the potential drop over the inner tube with that across a Tinsley 0.01 ohm standard resistance.

To determine  $R_0$  the resistance of the inner tube at 0 °C. and the temperature coefficient  $\alpha = \left( \frac{1}{R} \frac{dR}{dt} \right)_{0^\circ \text{C.}}$ , accurate linear extrapolations of the reciprocal of

the tube resistance  $1/R$  plotted against  $I^2$  to  $I^2=0$  were carried out at four different temperatures, namely, ice point, steam point, the transition point of pure sodium sulphate (32.38 °C.), and the carbon dioxide point (−78.50 °C.). The following values of  $R_0$  were obtained:

Ice point (0 °C.)	: 0.00766190 ohm
Steam point (100 °C.)	: 0.0115122 ohm
Sodium sulphate point (32.38 °C.)	: 0.0087992 ohm
Carbon dioxide point (−78.36 °C.)	: 0.0052311 ohm.

These points were fitted to the cubic formula

$$R_t = R_0(1 + at + bt^2 + ct^3),$$

the values of the constants being

$$a = 0.0044033; \quad b = 0.000005269; \quad c = 0.089496.$$

Accordingly at 0 °C. the temperature coefficient of the resistance has the value

$$\alpha = \left( \frac{1}{R} \frac{dR}{dt} \right)_{0^\circ \text{C.}} = 0.00440.$$

As in this work the measurements of the thermal conductivities of the gases investigated are restricted to the ice point, it was necessary to determine the

value of  $\lambda$  (given by (4)) at the ice point. For a hollow tube,  $\lambda$  is proportional to the true thermal conductivity  $\Omega$  of the tube, the connection between  $\lambda$  and  $\Omega$  being given by (5). The mean of six determinations of  $\lambda$  for a range of values of the current  $I$  was

$$\lambda = 0.01846 \pm 0.00006 \text{ cal. cm. sec.}^{-1} \text{ }^{\circ}\text{C.}^{-1}.$$

(The estimated true thermal conductivity of the tube given by (5) is

$$\Omega = 0.141 \text{ cal. cm.}^{-1} \text{ sec.}^{-1} \text{ }^{\circ}\text{C.}^{-1}.)$$

## VI. EXPERIMENTAL RESULTS

A careful comparison of the pressure dependence of the conductivity of two representative monatomic gases (argon and neon) and one diatomic gas (air) with that of the polyatomic gases carbon dioxide, nitrous oxide, and methane was made. The argon and neon from the British Oxygen Co., Wembley, were labelled spectroscopically pure. The sample of air used was dried by passing it over calcium chloride and then phosphorus pentoxide. The small amount of carbon dioxide present was not removed since its presence has no appreciable effect upon the observed conductivity.

The carbon dioxide was prepared in two ways: (a) by heating small fragments of magnesite in a hard glass tube and (b) by heating pure sodium bicarbonate. In both cases the gas was dried by passage over calcium chloride and then over phosphorus pentoxide. The results obtained by the two methods were completely concordant.

The nitrous oxide was obtained from a small cylinder of that gas supplied by Commonwealth Industrial Gases Ltd. Following the recommendation of Lord Rayleigh we condensed the cylinder contents by surrounding it with dry ice. About half the contents of the cylinder were then blown off, thus removing the only appreciable impurity—nitrogen.

The methane was obtained from the Old Balmain Colliery, N.S.W., through Tulloch Pty. Ltd., Sydney. The stated purity was 96 per cent. The principal impurities were air and smaller quantities of some ethylene derivatives. The gas as received was not further purified but was dried by passing it through calcium chloride and phosphorus pentoxide.

For the above gases except the argon and the neon, the apparatus was first thoroughly evacuated and then washed out several times with the gas before filling. The argon and neon were admitted into the evacuated apparatus from 1 litre gas globes sealed directly onto the apparatus.

In Table 1 the values of the conductivity  $k$  and the corresponding pressures are given for the various gases. In each case the conductivity has been corrected for radiation, for departure from strict radial flow, and reduced to 0  $^{\circ}\text{C.}$  by assuming a temperature coefficient of  $k$  equal to 0.003. The radiation amounts to an equivalent conductivity  $k_R$  of amount approximately  $0.038 \times 10^{-5}$  for each gas and this is subtracted from the calculated conductivity. To correct for departure from radial flow each value of  $k$  has been reduced by 1 per cent. The unit of thermal conductivity used throughout is the cal. cm. $^{-1}$  sec. $^{-1}$   $^{\circ}\text{C.}^{-1}$ ; the unit of the pressure  $p$  is the cm. of mercury. The temperatures quoted at the bottom of the columns are the mean temperatures of the gases.



The data given in Table 1 are shown graphed in Figures 2 and 3. The curves for the monatomic gases (argon and neon) and the diatomic gas (air) are completely flat over a wide range of the pressure. From the curve for air it is evident that the apparatus is certainly free from convection up to a pressure as high as 1 atm. The dip in the curves at the low pressure end merely indicates that the effect of temperature discontinuity is becoming appreciable. As neon has a high conductivity the effect of the temperature jump becomes noticeable at a higher pressure than is the case for air and argon. Summarizing, the apparatus is completely free from convection and the conductivity of the monatomic and the diatomic gases is independent of the pressure over a wide range of the pressure in accordance with Maxwell's prediction.

TABLE I  
THERMAL CONDUCTIVITY DATA

Monatomic Gases				Diatomic Gas		Polyatomic Gases					
Argon		Neon		Air		Carbon Dioxide		Nitrous Oxide		Methane	
<i>p</i>	$k \times 10^5$	<i>p</i>	$k \times 10^5$	<i>p</i>	$k \times 10^5$	<i>p</i>	$k \times 10^5$	<i>p</i>	$k \times 10^5$	<i>p</i>	$k \times 10^5$
				75.7	5.745	86.2	3.497	88.5	3.645	94.4	7.225
				65.1	5.748	74.5	3.487	71.5	3.631	85.2	7.225
				54.7	5.742	57.9	3.484	61.2	3.631	75.2	7.215
47.4	3.965			44.2	5.742	41.9	3.471	52.0	3.626	65.7	7.208
45.4	3.967	38.2	11.10	34.0	5.739	26.7	3.459	42.0	3.631	54.7	7.207
36.3	3.967	28.4	11.10	24.0	5.735	14.4	3.441	31.5	3.624	45.4	7.199
33.5	3.960	18.7	11.11	22.41	5.733	6.81	3.431	21.5	3.608	34.8	7.195
25.7	3.967	13.4	11.06	1.47	5.689	2.50	3.421	11.11	3.611	25.2	7.187
17.4	3.958	8.06	11.05			0.50	3.397	3.89	3.55	15.4	7.180
10.06	3.967	2.95	10.93								
4.33*	3.958	0.87	10.63							5.18	7.158
1.13	3.923									1.91	7.129
(2.0 °C.)		(0.9 °C.)		(1.4 °C.)		(1.9 °C.)		(1.8 °C.)		(1.2 °C.)	

In Figure 3 two of the gases, carbon dioxide and methane, show a small but appreciable decrease in  $k$  with the pressure over the whole range of pressure. We had no reason to doubt the reality of this effect. The experiments with carbon dioxide were repeated a number of times and two different methods of preparing the gas were used. The same slow decrease in  $k$  was always obtained. Since the apparatus is free from convection effects, and since the temperature discontinuity could have no appreciable effect at pressures greater than 20 cm. of mercury, the thermal conductivity of the two gases carbon dioxide and methane is not strictly independent of the pressure. Accordingly Maxwell's prediction from the classical kinetic theory does not apply for these two gases. Owing to the smallness of the effect at ordinary pressures, it would be difficult,

if not impossible, to detect it in any form of "hot wire" apparatus employing a fine filament. In an apparatus of that type the effect of the temperature discontinuity at the wire is appreciable at all pressures and would mask the effect observed in our apparatus.

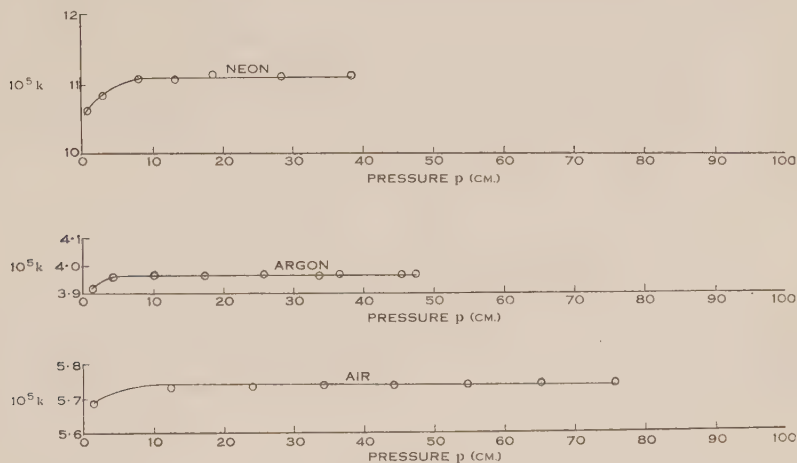


Fig. 2

The graph for nitrous oxide (Fig. 3) shows only an extremely small pressure dependence of  $k$ . The conductivity of this gas between 20 and 70 cm. of mercury is practically constant and has the value  $3.62 \times 10^{-5}$  cal. cm. $^{-1}$  sec. $^{-1}$  °C. $^{-1}$ .

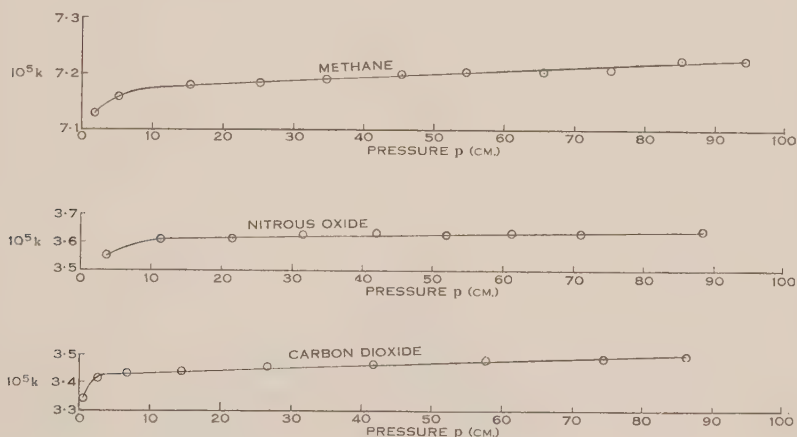


Fig. 3

Classical theory of the ratio of the two principal specific heats ( $\gamma$ ) leads to an expected value of 1.333 for all three gases. The experimental values are 1.300 for carbon dioxide, 1.313 for methane, and 1.324 for nitrous oxide. The agreement between theory and observation is much the closest for nitrous oxide.

It does not appear possible to account for the pressure dependence of  $k$  in the case of carbon dioxide and of methane observed in the present experiments except by postulating incomplete participation of the energy of vibration of the molecules of the two gases in the conduction process. One possible means of disclosing the existence of incomplete participation of the vibrational energy is to compare Maxwell's constant  $\varepsilon = \frac{k_0}{\eta_0 c_v}$ , where  $k_0$  and  $\eta_0$  are respectively the conductivity and viscosity at 0 °C. with Eucken's semi-empirical values of  $\varepsilon$ , namely,

$$\varepsilon_E = \frac{1}{4}(9\gamma - 5) \text{ for complete participation,}$$

and

$$\varepsilon_E = \frac{19}{4}(\gamma - 1) \text{ for no participation.}$$

TABLE 2

Gas	$k_0 \times 10^5$	$\eta_0 \times 10^6$	$c_v$	$\varepsilon = \frac{k_0}{\eta_0 c_v}$	$\gamma$	$\frac{1}{4}(9\gamma - 5)$	$\frac{19}{4}(\gamma - 1)$
CO <sub>2</sub>	3.43-3.49*	13.74	0.153	1.64-1.67	1.300	1.72	1.52
N <sub>2</sub> O	3.62	13.66	0.155	1.71	1.324	1.73	1.55

\* Between 8 and 80 cm. of mercury.

The data given in Table 2 are consistent with the pressure dependence of  $k$  observed in the present work. With nitrous oxide the close agreement between the values of  $\varepsilon$  and  $\frac{1}{4}(9\gamma - 5)$  indicates very nearly complete participation of the vibrational energy and accordingly  $k$  should be independent of the pressure. On the other hand the corresponding data for carbon dioxide would indicate incomplete participation and consequently an observable variation of  $k$  with the pressure.

The experiments described in this paper were not carried out primarily with the object of obtaining the most accurate absolute values of the conductivity of the gases investigated. The substitution of a nickel tube for the thick platinum wire employed in the experiments carried out in 1934 has the disadvantage that the electrical constants of nickel are less stable than those of platinum. There were, in fact, small irregular secular changes in the resistance of the nickel tube. However, the absolute values of  $k_0$  obtained in the present work agree well with the values previously obtained with the platinum wire except in the case of argon. The value of  $k_0 = 3.97 \times 10^{-5}$  is much higher than the value  $k_0 = 3.83 \times 10^{-5}$  obtained in 1934. The experiments with argon were repeated a number of times and merely serve to confirm the present result. A comparison of the values of  $k_0$  obtained in the present work with those obtained in 1934(2) is made in Table 3.

TABLE 3

Gas	Present Work	Kannuluik and Martin (1934)
Argon .. .. .	$3.97 \times 10^{-5}$	$3.83 \times 10^{-5}$
Neon .. .. .	11.08	11.12
Air .. .. .	5.74	5.76
Carbon dioxide .. ..	3.43-3.49	3.43
Nitrous oxide .. ..	3.61	3.61

## VII. ACKNOWLEDGMENTS

The authors wish to record their appreciation of the interest taken in the work by Professor L. H. Martin and to thank him for the facilities provided. The work formed part of a thesis submitted for the M.Sc. degree by one of the authors (H.B.D.).

## VIII. REFERENCES

- (1) UBBELOHDE, A. R.—*J. Chem. Phys.* **3**: 219 (1935).
- (2) KANNULUIK, W. G., and MARTIN, L. H.—*Proc. Roy. Soc. A* **144**: 496 (1934).
- (3) SCHÄFER, K., RATING, W., and EUCKEN, A.—*Ann. Phys.* **42**: 176 (1942).
- (4) WEBER, S.—*Ann. Phys.* **54**: 326 (1917).
- (5) ULSAMER, J.—*Z. Ver. deutsch. Ing.* **80**: 537 (1936).
- (6) KANNULUIK, W. G., and LAW, P. G.—*Proc. Roy. Soc. Vict.* **58**: 142 (1947).
- (7) KANNULUIK, W. G., and MARTIN, L. H.—*Proc. Roy. Soc. A* **141**: 144 (1933).



# ASYMPTOTIC EXPRESSIONS FOR THE ENERGIES OF CERTAIN LONG MOLECULES

By R. D. BROWN\*

[Manuscript received May 11, 1950]

## Summary

Asymptotic values for the increment in total  $\pi$ -electron energy and in resonance energy for addition of one structural unit to a long polyene and to a long polyacene molecule are derived. The values are obtained, both neglecting and including the overlap integral, by the usual LCAO molecular-orbital method. The results are compared with the corresponding increments for the lower members of each series of molecules calculated by direct summation.

## I. INTRODUCTION

In the quantum-mechanical treatment of molecular structure by the LCAO molecular-orbital method, solution of the secular determinant provides a series of roots,  $x_r$ , the total energy of the electrons thus considered being given by (1)

$$E = 2n\alpha - 2 \sum_{r=1}^n x_r \beta, \quad \dots\dots\dots (1)$$

when there are  $2n$   $\pi$ -electrons and the molecule is in its ground state. Formula (1) involves the assumption that the overlap integral between neighbouring atoms is zero whereas it is known (1) to have a value close to 0.25. Inclusion of this integral in the calculations leads to the formula (1)

$$E' = 2n\alpha - 2 \sum_{r=1}^n e_r \gamma, \quad \dots\dots\dots (2)$$

where  $e_r$  is related to  $x_r$  by the formula

$$e_r = \frac{x_r}{1 - x_r S}, \quad \dots\dots\dots (3)$$

$S$  being the overlap integral.

For several types of long molecule it has proved possible to derive general expressions for  $x_r$ , notably for the linear polyenes(2), and for a series of linearly fused benzene rings(3). In certain applications of theoretical results to predict chemical properties(1, 4), it is useful to have asymptotic values for the increase in total energy or in the resonance energy of such molecules caused by the addition of one more of the repeating units. The purpose of the present paper is to communicate such results obtained for the polyenes and the polyacenes by a simple, general method of calculation.

\* Chemistry Department, University of Melbourne.

## II. CALCULATIONS

For the polyene  $C_m H_{m+2}$  the expression for  $x_r$  is (2)

$$x_r = -2 \cos r\pi/(m+1), \quad \dots\dots\dots (4)$$

the orbitals occupied in the ground state corresponding to all values of  $r$  ranging from 1 to  $m/2$ , so that  $m/2$  is to be substituted for  $n$  in (1). The method of obtaining the asymptotic expression for (1) consists in replacing the summation by the integral to which it tends. Thus for the linear polyenes we have

$$\begin{aligned} E_m &= m\alpha - 2\beta \sum_{r=1}^{m/2} x_r \\ &= m\alpha + 4\beta \sum_{r=1}^{m/2} \cos r\pi/(m+1) \\ &\sim m\alpha + 4\beta \frac{m+1}{\pi} \lim_{m \rightarrow \infty} \sum_{r=1}^{m/2} \frac{\pi}{m+1} \cos \left( \frac{\pi}{m+1} \right) \\ &= m\alpha + 4\beta \frac{m+1}{\pi} \int_0^{\pi/2} \cos \psi d\psi \\ &= m\alpha + 4(m+1)\beta/\pi. \end{aligned}$$

The next member of the series is  $C_{m+2} H_{m+4}$  so that

$$\delta E \sim 2\alpha + 8\beta/\pi \approx 2\alpha + 2 \cdot 546479\beta.$$

The resonance energy,  $R$ , of a molecule may be defined(1) as the difference (with positive sign) between the total  $\pi$ -electron energy and the total energy of the same number of ethylenic  $\pi$ -electrons. Thus we derive

$$\delta R = 2(\alpha + \beta) - \delta E \sim -0 \cdot 546479\beta. \quad \dots\dots\dots (5)$$

Next, substituting (4) in (3), we obtain in the same way the corresponding asymptotic expression for  $E'_m$ , assuming  $S$  to be exactly 0.25

$$E'_m \sim m\alpha + \frac{8(m+1)\gamma}{\pi} \int_0^{\pi/2} \frac{\cos \psi d\psi}{2 + \cos \psi},$$

the integration being performed using the standard substitution  $\tan \psi/2 = u$  to give

$$\begin{aligned} E'_m &\sim m\alpha + \left( 4 - \frac{16}{3\sqrt{3}} \right) (m+1)\gamma, \\ &\approx m\alpha + 0 \cdot 92080(m+1)\gamma, \end{aligned}$$

and from this

$$\delta E' \sim 2\alpha + 1 \cdot 84160\gamma,$$

similarly

$$\delta R = 2\alpha + 1 \cdot 6\gamma - \delta E' \sim -0 \cdot 24160\gamma. \quad \dots\dots\dots (6)$$

Turning now to the polyacenes, Coulson(3) has shown that for the polyacene  $C_{4m+2} H_{2m+4}$  the roots of the secular determinant are given by

$$\begin{aligned} (a) & \quad -(\sqrt{9+8 \cos r\pi/(m+1)}+1)/2, \\ (b) & \quad -(\sqrt{9+8 \cos r\pi/(m+1)}-1)/2, \end{aligned} \quad \dots\dots\dots (7)$$

those corresponding to levels occupied in the ground state corresponding to values of  $r$  ranging from 1 to  $m+1$  in (a) and 1 to  $m$  in (b). It is easy to see that

the summation over the  $2m+1$  levels occupied in the ground state required for (1), may be written as

$$\Sigma x_r = - \sum_{r=1}^{m+1} \sqrt{9+8 \cos r\pi/(m+1)}, \quad \dots\dots\dots (8)$$

the inclusion of  $2m+2$  terms from (7) being justified because the additional term is zero. Using (8) we find

$$E_m \sim (4m+2)\alpha + \frac{6(m+1)\beta}{\pi} \int_{\pi/(m+1)}^{\pi} \frac{1+p \cos \psi d\psi}{\pi/(m+1)},$$

where  $p$  represents  $8/9$  and  $\psi = \pi r/(m+1)$ . The lower terminal of the integral of course tends to zero and the further substitution  $\psi/2 = u$  transforms the expression to

$$(4m+2)\alpha + \frac{12(m+1)(1+p)^{\frac{1}{2}}}{\pi} \beta \int_0^{\pi/2} (1-k^2 \sin^2 u)^{\frac{1}{2}} du,$$

where  $k^2 = 2p/(1+p)$ . The integral is now in the form of the complete elliptic integral of the second kind so we have

$$\begin{aligned} E_m &\sim (4m+2)\alpha + \frac{12(m+1)(1+p)^{\frac{1}{2}}}{\pi} E(k)\beta, \\ &= (4m+2)\alpha + \frac{4\sqrt{17}(m+1)}{\pi} E\left(\sqrt{\frac{16}{17}}\right)\beta. \end{aligned}$$

Substituting the appropriate numerical values, the increase in energy caused by addition of another ring to give  $m+1$  rings in all consequently causes an increase in energy approaching the limit

$$\delta E_m = 4\alpha + 5.6113508\beta,$$

the corresponding limiting value for the increase in resonance energy being

$$\delta R = -1.6113508\beta. \quad \dots\dots\dots (9)$$

Next we may calculate  $\delta R'$ . Substituting the expressions (7) in (3) we obtain

$$\begin{aligned} (a) \quad & -4 \left\{ \frac{(9+8 \cos \psi)^{\frac{1}{2}} - \cos \psi}{9 - \cos \psi} \right\}, \\ (b) \quad & -4 \left\{ \frac{(9+8 \cos \psi)^{\frac{1}{2}} - 2 - \cos \psi}{5 - \cos \psi} \right\}, \end{aligned} \quad \dots\dots\dots (10)$$

and the limiting expression for  $\sum_1^{2m+1} e_r$  thus tends to

$$\begin{aligned} & -\frac{4(m+1)}{\pi} \left\{ \int_0^{\pi} \frac{(9+8 \cos \psi)^{\frac{1}{2}} d\psi}{9 - \cos \psi} - \int_0^{\pi} \frac{\cos \psi d\psi}{9 - \cos \psi} \right. \\ & \quad \left. + \int_0^{\pi} \frac{(9+8 \cos \psi)^{\frac{1}{2}} d\psi}{5 - \cos \psi} - \int_0^{\pi} \frac{2 + \cos \psi d\psi}{5 - \cos \psi} \right\}. \end{aligned}$$

The second and fourth integrals may be evaluated by standard methods to give

$$\sum_1^{2m+1} e_r \sim -\frac{4(m+1)}{\pi} \left\{ I_1 - \pi \left( \frac{9}{4\sqrt{5}} - 1 \right) + I_2 - \pi \left( \frac{7}{2\sqrt{6}} - 1 \right) \right\}. \quad \dots\dots\dots (11)$$

The integrals  $I_1$  and  $I_2$  are both of the form

$$I = \int_0^\pi \frac{(1+p \cos \psi)^{\frac{1}{2}} d\psi}{1-\alpha \cos \psi}.$$

This may be expressed in terms of elliptic integrals by appropriate substitutions and rearrangements.\* However, as tables of elliptic integrals of the third kind were not available the integrals  $I_1$  and  $I_2$  in (11) were evaluated numerically using Weddle's 6-ordinate method with the sixth difference correction(5) over each of the two quadrants. The results were

$$I_1 = 0.32286\pi,$$

$$I_2 = 0.60289\pi.$$

Substituting these and the other numerical values in (11) we obtain the value  $0.49065\pi$  for the expression inside the braces so that, using (2),

$$E'_m \sim (4m+2)\alpha + 3.92521(m+1)\gamma.$$

Thus in the limit

$$\begin{aligned} \delta E' &= 4\alpha + 3.92521\gamma \\ \delta R' &= -0.72521\gamma \end{aligned} \quad \dots\dots\dots (12)$$

The results (5), (6), (9), and (12) may now be compared with the increments for the first few members of each series of molecules obtained by direct summation of the appropriate terms.

### III. DISCUSSION OF RESULTS

The increments in  $R$  and  $R'$  for the first six polyenes are listed in Table 1. Clearly the convergence is fairly slow, the increments last listed still differing from the limiting values even in the second decimal place.

TABLE 1  
THE POLYENES

	$R$	$\delta R$	$R'$	$\delta R'$
Ethylene ..	0		0	
Butadiene ..	0.4721	0.4721	0.1747	0.1747
Hexatriene ..	0.9879	0.5158	0.3868	0.2121
Octatetraene	1.5175	0.5296	0.6119	0.2251
Decapentaene	2.0533	0.5358	0.8430	0.2311
Dodecahexaene	2.5925	0.5392	1.0673	0.2243

\*  $I = 2 \frac{(1+p)^{\frac{1}{2}}}{1-\alpha} \left\{ \frac{u+k^2}{u} \Pi(k, u) - \frac{k^2}{u} F(k) \right\}$ , where  $k^2 = 2p/(1+p)$ ,  $u = 2\alpha/(1-\alpha)$ , and the functions

$F$  and  $\Pi$  are complete elliptic integrals of the first and third kinds respectively.



The corresponding increments for the first six polyacenes are given in Table 2. In this case the convergence is much more rapid, the last values listed in the table for  $\delta R$  and  $\delta R'$  differing from the limiting values only by about 1 part in 2000 and 1 part in 1000 respectively. This is doubtless partly connected with the fact that the summation is over a greater number of energy levels for the polyacene than for the corresponding polyene.

TABLE 2  
THE POLYACENES

	$R$	$\delta R$	$R'$	$\delta R'$
Benzene ..	2	1.6832	1.0667	0.7964
Naphthalene ..	3.6832	1.6305	1.8631	0.7446
Anthracene ..	5.3137	1.6171	2.6077	0.7312
Naphthacene	6.9308	1.6132	3.3389	0.7271
Pentacene ..	8.5440	1.6120	4.0660	0.7260
Hexacene ..	10.1560		4.7920	

These results have already been employed for the calculation of *para* localization energies(4, 6), and also bond localization energies(1, 7).

#### IV. ACKNOWLEDGMENT

The helpful suggestions made by Dr. E. R. Love, Mathematics Department, University of Melbourne, after reading the initial draft of this paper are very gratefully acknowledged.

#### V. REFERENCES

- (1) BROWN, R. D.—*Aust. J. Sci. Res. A* **2**: 564 (1949).
- (2) COULSON, C. A.—*Proc. Roy. Soc. A* **169**: 413 (1939).
- (3) COULSON, C. A.—*Proc. Phys. Soc. A* **40**: 257 (1948).
- (4) BROWN, R. D.—*J. Chem. Soc.* **1950**: 691 (1950).
- (5) COMRIE, L. J.—“Chambers's Six-Figure Mathematical Tables.” Vol. 2, p. 547 (1949).
- (6) BROWN, R. D.—*J. Chem. Soc.* **1950**: 2730 (1950).
- (7) BROWN, R. D.—*J. Chem. Soc.* (in press).

# A STUDY OF THE OXIDATION OF ETHYLENE TO ETHYLENE OXIDE ON A SILVER CATALYST

By K. E. MURRAY\*

[*Manuscript received March 30, 1950*]

## *Summary*

A study has been made of the direct oxidation of ethylene to ethylene oxide in a flowing system under the catalytic action of silver. Various catalysts were prepared and one containing 10 per cent. barium carbonate, having a high and sustained activity, was selected for a detailed examination of the influence of reaction conditions on the yield of ethylene oxide.

It has been found that the percentage ethylene in the reaction mixture can be varied widely without greatly changing the yield of ethylene oxide, but that the conversion is decreased by increasing the ethylene content. The yield increases with increasing oxygen content until the latter reaches 20 per cent. after which further enrichment with oxygen has little effect. The presence of water vapour slightly increases the yield, but considerably decreases conversion. The yield is much decreased by diminishing the reaction pressure. Highest yields (73 per cent.) were obtained when small amounts of ethylene dichloride were added to the reactants but at the expense of low conversion. Propylene and *isobutylene* yielded only carbon dioxide and water when oxidized in the presence of a silver catalyst.

The results are discussed in relation to the reaction mechanism suggested by Twigg(1).

## I. INTRODUCTION

This investigation was made during the years 1943-45. When first begun the only information on the catalytic oxidation of ethylene to ethylene oxide was contained in the patent literature. The object was to prepare an active stable catalyst and with it to examine the influence of variables such as temperature, pressure, and gas composition upon the reaction in a flowing system.

Since that time accounts of other investigations of this reaction have appeared in the scientific literature(2-6). Knowledge of the reaction mechanism has been increased considerably by the comprehensive investigations of Twigg(1, 7). As a result of his work it is known that on the silver catalyst some of the ethylene is oxidized to carbon dioxide by way of very unstable intermediates. The remainder of the ethylene oxidized is converted first to ethylene oxide some of which is then isomerized on the catalyst to acetaldehyde which undergoes rapid oxidation to carbon dioxide and water. On the publication of Twigg's findings further proposed work in this Laboratory on the reaction mechanism was considered unjustified. However, as particular care had been taken in the design of apparatus to ensure good control of the reaction conditions, and since the conditions relate more closely to the industrial process (e.g. high space velocities) than those used by Twigg, the results are presented, and are discussed in relation to his findings on the reaction mechanism.

\* Division of Industrial Chemistry, C.S.I.R.O., Melbourne.

## II. APPARATUS AND METHOD

Figure 1 shows diagrammatically the arrangement of the apparatus used in the greater part of this work. It was above laboratory size and was designed originally to test catalysts for activity and useful life, particular attention being paid to regulation of temperature and flow-rate in order to obtain good control of reaction conditions with the minimum of attention.

The reaction chamber unit consisted of a copper tube *A*, 4 ft. by  $\frac{7}{8}$  in. internal diameter, silver plated internally, and surrounded by a mild steel jacket through which liquid "Dowtherm" was circulated by a small stainless steel centrifugal pump *B*, fitted with a water-cooled gland. The "Dowtherm" also circulated through leg *C* of the unit, which was wound with a heating element regulated by a temperature controller operating from a thermocouple in a pocket at point *D*. The jacket around *A* was also electrically wound to compensate for heat losses and give a more sensitive operation of the controller. The temperature given by a thermometer placed in a pocket near *D* was accepted as the reaction temperature. The temperature "hunted" evenly over the range given in the tables (usually 3–4 °C.): for purposes of calculation the average temperature of the range was used.

In order to obtain good heat transfer from the catalyst and reproducible flow conditions in the reaction zone, the catalyst was supported as a thin film on metal fins as suggested by Cambron *et al.*(2). These were made in sections 10 cm. in length and fitted closely the inside of the reaction tube *A*. As shown in the insert of Figure 1 each section was made in two halves, the fins of one half interleaving the fins of the other. A piece of bent spring steel *E* held the two halves in close contact with the wall of the reaction tube. They were chrome plated to permit easy removal of a tested catalyst with nitric acid. Four sets of fin supports were used in the lower part of the reaction tube, the upper part functioning as a preheater for the gases. To determine whether this method of mounting the catalyst was efficient in removing the heat of reaction, a bare thermocouple was mounted in contact with the central fin and the reading during a preliminary run with an active catalyst at 260 °C. compared with the thermometer in the heat transfer liquid (at *D*). The difference was within the estimated overall accuracy of measurement (3–4 °C.).

The compositions of the reaction mixtures were governed by controlling the flow-rates of ethylene and air and mixing the gases just before the reaction chamber. Air was supplied by a compressor at 10–15 p.s.i. gauge to a simple photoelectric "on-off" controller *F*. A tank *H* was included to damp out surges. Ethylene of medical grade (99.5–99.6%, Ohio Chemical Co.) was contained in a bell-type gas holder *K* under slight pressure and its rate of flow was adjusted by a needle valve and read on a metal-glass flowmeter *L*. At times other gases, N<sub>2</sub>, O<sub>2</sub>, CO<sub>2</sub>, propylene, and isobutylene were used and their flow-rates were controlled either by the controller mentioned above or by an arrangement similar to that for ethylene.

The experiments carried out under reduced pressure were made with a reaction chamber consisting of a stainless steel tube ( $\frac{7}{8}$  in. diameter, 15 in.

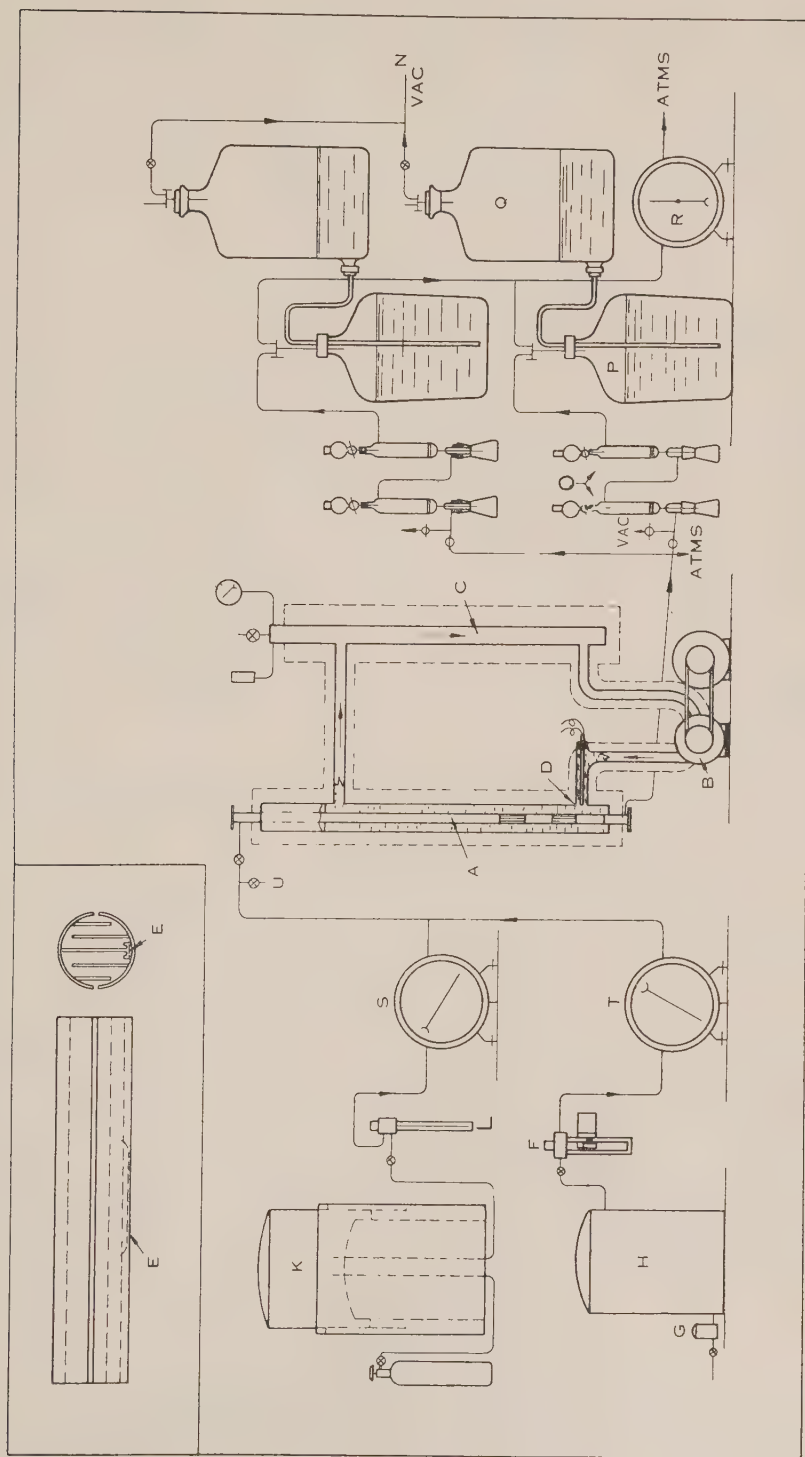


Fig. 1.—Diagram of apparatus. Insert—A catalyst fin support.



long) surrounded by a metal jacket containing a heat transfer liquid (dimethyl phthalate, b.p. 283 °C.) which was boiled by an external electrical winding, the temperature of the boiling liquid being regulated by applying a slightly reduced constant pressure above the condenser by means of a manostat. The pressure on the reaction system was controlled by a second manostat which, after the removal of excess ethylene, ethylene oxide,  $\text{CO}_2$ , and  $\text{H}_2\text{O}$  by a liquid air trap, vented the excess oxygen into the line vacuum.

### III. METHODS OF ANALYSIS

Ethylene oxide in the reacted gases was estimated by adsorption in saturated aqueous calcium chloride containing approximately 0.5N HCl. This reagent described by Kerchov(8) was found to be as effective and cheaper than the magnesium chloride reagent of Lubatti(9). By the application of a vacuum at *N* (Fig. 1) a portion of the reacted gases was drawn through two gas washers *O* in series each containing 10 ml. of the reagent. Using gas washers with sintered glass plates of medium porosity, and a flow-rate of 12–24 l./hr., adsorption of oxide was almost complete in the first washer. The residual gas containing  $\text{CO}_2$  was collected over brine in the vessel *P*. Its volume was measured at the end of the analysis by opening vessel *Q* to the atmosphere, allowing the brine to fill *P* and to displace the gas through wet test meter *R*. A small amount (1–2%) of the HCl is carried off from the reagent by the gas: the correction to be applied was determined by passing an equal volume of air, usually 6 to 10 l., through the gas washers. Residual HCl was titrated by standard NaOH using methyl red as the indicator.

The analysis of  $\text{CO}_2$  was carried out simultaneously and in like apparatus, the  $\text{CO}_2$  being absorbed from another portion of the reacted gases in approximately 0.4N NaOH, excess NaOH being titrated with HCl after precipitating carbonate with barium chloride.

In calculating percentage oxide and  $\text{CO}_2$  in the reacted gases, corrections were made to the residual gases for the removal of water and ethylene oxide in the analysis for oxide, and water, oxide, and  $\text{CO}_2$  in the analysis for  $\text{CO}_2$ . When conversions were large, correction was also made for the contraction in volume which accompanies the formation of oxide. Analyses could be made without interrupting the operation of the apparatus. The ethylene content of the feed gases was calculated from simultaneous readings of the wet test meters *S* and *T* at the beginning and end of each analysis, which usually occupied 10–20 minutes.

From these analyses the conversions per pass of the ethylene in the reactant mixture to ethylene oxide and to  $\text{CO}_2$  were calculated and since these are the only carbon compounds the yield of ethylene oxide based on the ethylene reacted was obtained from the expression

$$\frac{\% \text{ conversion to oxide}}{\% \text{ conversion to oxide} + \text{CO}_2} \times \frac{100}{1}.$$

### IV. PREPARATION OF THE CATALYST

The catalyst for detailed examination was chosen after many catalysts had been prepared and their activities compared. Some of these preparations

contained additives such as Ca, Sr, Ba, Mg, carbonates, silica, boric acid, etc. to determine whether they functioned as "promoters" in enhancing catalytic activity. As a result it appeared that the function of such a second component in the catalyst was to preserve its original activity by preventing or reducing sintering. It was found that pure silver catalysts lost their activity quickly. For instance, a sample of finely divided silver prepared by the electrolysis of a silver nitrate solution under a high current density had an initial activity comparable to the most active catalysts tested; but in 20 hours' use at 260 °C. half this activity had been lost. Electron micrographs were taken before and after use, and these showed that the catalyst was originally in the form of interlaced and branching rods, which after use lost their fine structure and appeared more massive. The general appearance of the catalyst at the same time had changed from dark grey to light grey.

In preparing an active and stable catalyst it was considered essential to obtain the silver in a finely divided state intimately mixed with the second component. The method adopted was co-precipitation of a mixture of silver and barium carbonates from a solution of the nitrates, followed by reduction of the washed and dried precipitate to silver with hydrogen. The reduction is exothermic and much sintering occurs unless the heat liberated is adequately dissipated. This was achieved by reducing the mixed carbonates after they had been applied to the metal carriers.

The catalyst used in this work was prepared as follows. Silver nitrate (A.R., 39.35 g.) and barium nitrate (A.R., 3.35 g.), were dissolved in 500 ml. of distilled water. Sodium carbonate (A.R., 14 g.) was dissolved in a second 500 ml. of distilled water and added with stirring to the first solution. The precipitate was allowed to settle, washed three times by decantation, filtered, and dried in vacuum. Carbonates (15 g.) was milled with 7.5 ml. of ethylene glycol in a rod mill for 4 hours and the resulting thin cream applied evenly to four sets of catalyst fin supports, the glycol removed under vacuum at 100 °C., and the silver carbonate reduced at this temperature by the passage of hydrogen for 30 min. The catalyst was dark grey in appearance. The weight of catalyst on the carriers was 6.85 g. distributed on a calculated surface of 800 sq. cm. in a reaction volume of 90 cc.

This preparation at 240 °C. gave an initial yield of 45.0 per cent. oxide. Continued use at this temperature for 3 working days (estimated total time 24 hours) caused reduction of between 10 and 20 per cent. in conversions to oxide and CO<sub>2</sub>, the latter being slightly greater and causing an increase of yield of oxide to 47 per cent. In order to bring the catalyst to a steady state more rapidly the temperature was raised to 270 °C. and nine identical runs were made, by which time the catalyst had become stable. Space velocity\* was 2800  $\pm$  100 and ethylene content of the reaction mixture in the limits 4.34–5.19 per cent.

\* The term space velocity as used is defined as the ratio  
$$\frac{\text{volume of gases passed per hour at N.T.P.}}{\text{reaction volume}}$$

The lower term includes the catalyst and the space between the catalyst carriers but excludes the volume of the carriers.

The results are plotted in Figure 2 which shows clearly the rapid approach to equilibrium of the catalyst. There is an initial increase in yield of oxide due to an increase in conversion to oxide and a correspondingly greater decrease in conversion to  $\text{CO}_2$ , the total conversion being practically unchanged. The time taken for this approach to equilibrium (the first five runs) was estimated to be between 8 and 12 hours.

The improvement in yield on use was also observed with another catalyst of comparable activity, prepared in the same way, but substituting calcium

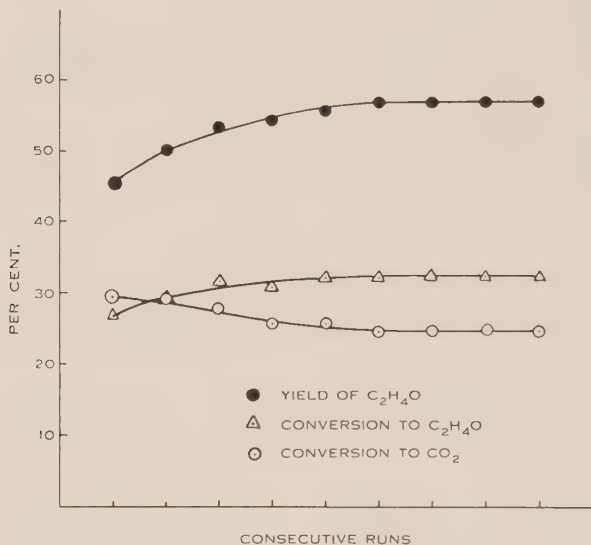


Fig. 2.—Approach to equilibrium of catalyst.

carbonate for barium carbonate. These slow changes in activity demonstrate the impossibility of adequately comparing the activity of catalysts for this reaction without their use for considerable periods.

An electromicrograph of the silver-barium carbonate catalyst was made when the investigation was complete. It appeared to be made up of indefinite crystalline particles with the pronounced rod formation of the electrolytic sample absent. Definition was not good enough to differentiate between particles of silver and barium carbonate.

## V. RESULTS

### (a) *The Influence of Water Vapour on the Reaction*

Three runs have been made under similar conditions with the exception that, in the second run the water vapour was largely removed by passing the gases through two traps cooled by dry ice-acetone to between  $-70$  and  $-80^\circ\text{C}$ . In the other runs the gases were presumably saturated with water vapour at room temperature because they had passed over water in wet-test meters. The results are summarized in Table 1.

TABLE I  
 THE EFFECT OF DRYING THE REACTANT GASES

Run No.	Condition of Gases	Temp. (°C.)	C <sub>2</sub> H <sub>4</sub> in Reaction Mixture (%)	Conversions		Yield Oxide (%)
				To Oxide (%)	To CO <sub>2</sub> (%)	
68a	Gases not dried ..	248-251	4.64	18.9	11.2	62.8
b		„	4.60	20.0	12.2	62.2
69a	Gases dried by cooling to -70 to -80 °C.	248-251	4.55	25.9	16.6	60.9
b		„	4.69	25.5	16.6	60.5
70a	Gases not dried ..	248-251	4.64	21.0	12.9	61.8
b		„	4.73	20.6	12.9	61.7

By averaging the figures from runs 68 and 69 it is found that removal of water vapour has increased the conversion to oxide by 31 per cent. At the same time there occurs a small but significant decrease in the yield of oxide. In all subsequent runs the gases have not been dried. After being stored over water and passed over water in wet-test meters they were considered to be in a standard condition for the examination of other variables.

(b) *The Ethylene Content of the Gases*

In varying the ethylene content of the reactant gases the actual variable involved is the air-ethylene ratio, but the results are clearer if expressed as per cent. ethylene in the mixture, bearing in mind the fact that the oxygen content also varies slightly over the range examined. Table 2 summarizes the results from two series of runs, the second series (Nos. 72-82) being carried out three days after the first series.

The above results are plotted in Figure 3 as per cent. ethylene in the reaction mixture against conversions to oxide and to CO<sub>2</sub>, yield of oxide, and volume (l./hr.) of ethylene oxidized to oxide. It is found on increasing the ethylene content that (i) the conversions to oxide and to CO<sub>2</sub> decreased over the whole range, (ii) the yield of oxide decreased slightly, and (iii) the volume of ethylene oxidized to oxide increased up to a maximum at 12-13 per cent. ethylene above which it decreased. The maximum for the second series of runs occurred at 15 per cent. ethylene. Owing to the small change in yield the volume oxidized to CO<sub>2</sub> follows the same trend as the volume oxidized to oxide.

(c) *The Oxygen Content of the Gases*

The oxygen content of the gas mixture was varied by mixing oxygen, nitrogen, and ethylene. The ethylene content, space velocity, and temperature were kept constant. Table 3 summarizes the results of six runs covering the



TABLE 2  
 THE EFFECT OF ETHYLENE CONTENT OF THE GASES ON THE REACTION

Run No.	Temp. (°C.)	C <sub>2</sub> H <sub>4</sub> in Reaction Mixture (%)	Air/C <sub>2</sub> H <sub>4</sub> Ratio	Space Velocity	C <sub>2</sub> H <sub>4</sub> Converted to Oxide (l./hr.)	Conversions		Yield Oxide (%)
						To Oxide (%)	To CO <sub>2</sub> (%)	
71a	248-251	1.88	52.0 : 1	1490	0.69	27.1	16.7	61.9
71b	"	2.10	46.7 : 1	1450	0.74	27.6	17.0	61.9
72a	"	4.39	21.7 : 1	1490	1.19	20.4	13.1	60.9
72b	"	4.39	21.7 : 1	1490	1.26	21.3	13.3	61.6
73	"	6.25	15.0 : 1	1480	1.53	18.4	12.1	60.3
74	"	8.26	11.1 : 1	1500	1.77	15.8	10.5	60.2
75	"	9.97	9.2 : 1	1490	1.89	14.4	9.9	59.2
76	"	11.77	7.5 : 1	1490	2.07	13.2	9.5	58.0
77	"	14.28	6.0 : 1	1490	1.98	10.4	7.7	57.8
79	247-250	17.72	4.6 : 1	1520	1.97	8.1	5.8	58.4
80	"	14.65	5.8 : 1	1500	2.11	10.7	7.6	58.3
81	"	11.91	7.4 : 1	1510	1.97	12.2	8.6	58.7
82	"	9.8	9.2 : 1	1500	1.83	13.9	9.9	58.4

range 23.1 down to 2.4 per cent. oxygen, and Figure 4 shows these results plotted as conversions and yield of oxide against oxygen content.

 TABLE 3  
 THE EFFECT OF OXYGEN CONTENT OF THE GASES ON THE REACTION

Run No.	Temp. (°C.)	O <sub>2</sub> in Reaction Mixture (%)	C <sub>2</sub> H <sub>4</sub> in Reaction Mixture (%)	Space Velocity	Conversions		Yield Oxide (%)
					To Oxide (%)	To CO <sub>2</sub> (%)	
123	262-265	23.1	4.61	1500	31.0	20.3	60.4
124	"	15.6	4.53	"	25.7	18.1	58.6
125	"	10.4	4.60	"	21.2	15.6	57.6
126	"	7.54	4.63	1490	17.3	15.1	53.4
127	"	5.00	4.57	"	12.6	13.6	48.1
128	"	2.40	4.64	"	7.1	10.5	40.2

A decrease in oxygen content below that in air decreased both the conversion and the yield of oxide. A corresponding increase might be expected by carrying out the reaction in oxygen. In two comparable runs at 247-250 °C. with 4.68 per cent. ethylene in air and 4.77 per cent. ethylene in oxygen the overall conversion was increased by 31 per cent. with but a slight increase in the yield of oxide from 58.0 to 60.5 per cent.

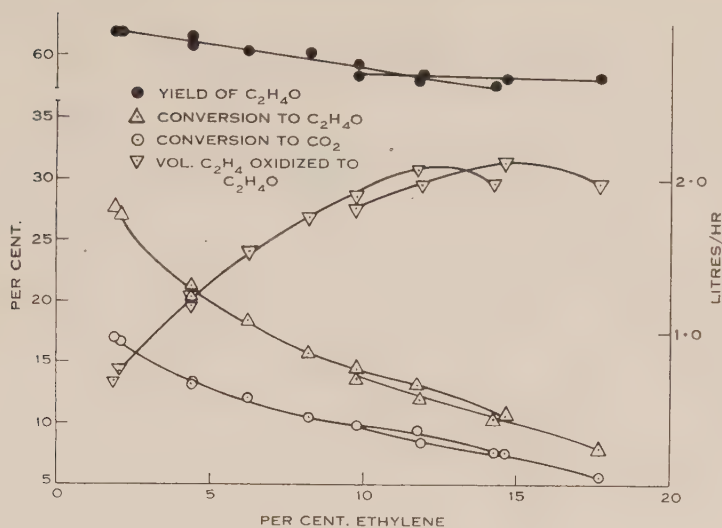


Fig. 3.—Effect of ethylene content on the reaction.

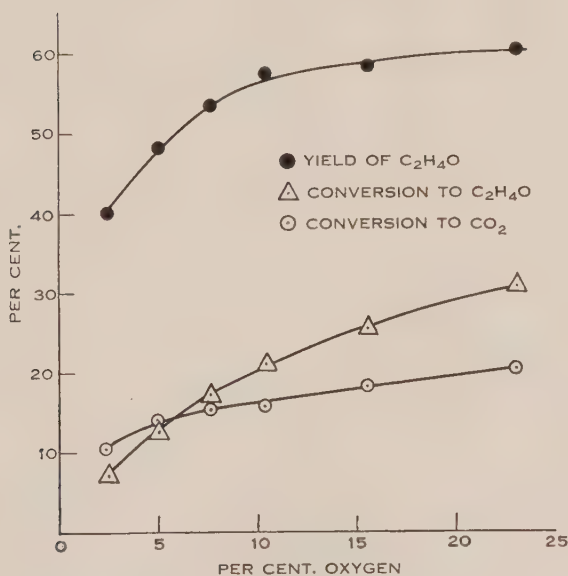


Fig. 4.—Effect of oxygen content on the reaction.

*(d) Space Velocity*

The results of nine consecutive runs, increasing space velocity over the range 600 to 4060, other factors being kept constant, are summarized in Table 4 and presented graphically in Figure 5.

Increasing space velocity decreases the conversion, increases the reaction rate, and affects the yield of oxide but little. For this reason the amount of ethylene oxide produced increases with space velocity. For a 6.8-fold increase in space velocity the conversion to oxide was decreased by a factor of 3.3.

TABLE 4  
 EFFECT OF SPACE VELOCITY ON THE REACTION

Run No.	Temp. (°C.)	C <sub>2</sub> H <sub>4</sub> in Reaction Mixture (%)	Space Velocity	Oxidized to Oxide (l./hr.)	Conversions		Yield Oxide (%)
					To Oxide (%)	To CO <sub>2</sub> (%)	
83	247-249	4.93	600	0.85	35.1	22.1	61.4
84	"	4.74	730	0.96	31.8	20.3	61.1
85	"	4.72	980	1.12	27.4	17.5	61.0
86	"	4.99	1190	1.29	24.6	16.4	60.0
87	"	4.69	1550	1.31	20.0	14.1	60.1
88	"	4.69	2070	1.50	17.4	11.9	59.4
89	"	4.54	2780	1.60	14.9	10.1	59.6
90	"	4.75	3290	1.74	12.5	8.8	58.7
91	"	4.73	4060	1.82	10.6	7.8	57.6

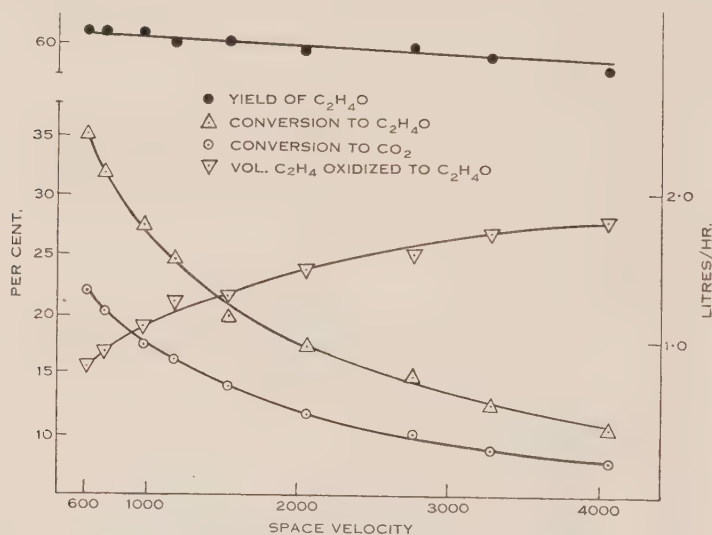


Fig. 5.—Effect of space velocity on the reaction.

## (e) Temperature

The effect of temperature has been studied over the range 224 to 283 °C. at a space velocity of 1480-1510 and from 274 to 315 °C at a space velocity 3700-3790. The reaction conditions and results of 15 runs are set out in Table 5 and plotted in Figure 6. Runs 108-115, 116-120, and 121-122 were carried out on consecutive days, the reaction chamber being allowed to cool overnight.

Up to 250 °C. an increase of temperature increased both conversion to oxide and to CO<sub>2</sub> without affecting the yield of oxide. Above 250 °C. the yield slowly

TABLE 5  
 THE EFFECT OF TEMPERATURE

Run No.	Temp. (°C.)	C <sub>2</sub> H <sub>4</sub> in Reaction Mixture (%)	Space Velocity	Conversions		Yield Oxide (%)
				To Oxide (%)	To CO <sub>2</sub> (%)	
108	224 ±1	4.45	1510	10.9	7.2	60.0
109	"	4.78	1500	9.7	6.6	59.5
110	245 ±1	4.81	1500	19.0	12.9	59.6
111	"	4.79	1500	19.0	12.9	59.6
112	266 ±1	4.76	1490	29.6	21.5	57.9
113	266.5 ±1	4.77	1490	30.3	22.4	57.5
114	283 ±1.5	4.77	1480	37.5	31.9	54.0
115	"	4.84	1490	36.8	31.0	54.2
116	274 ±1	4.73	3730	18.7	15.6	54.5
117	287 ±1.5	4.74	3770	23.6	21.2	52.7
118	303 ±2	4.84	3700	28.9	28.8	50.2
119	315 ±1.5	4.84	3770	30.7	35.8	46.2
120	272 ±1	4.77	3790	18.5	18.7	49.7
121	274.5 ±1	4.87	3760	20.0	18.3	52.2
122	"	4.84	3730	18.1	15.9	53.2

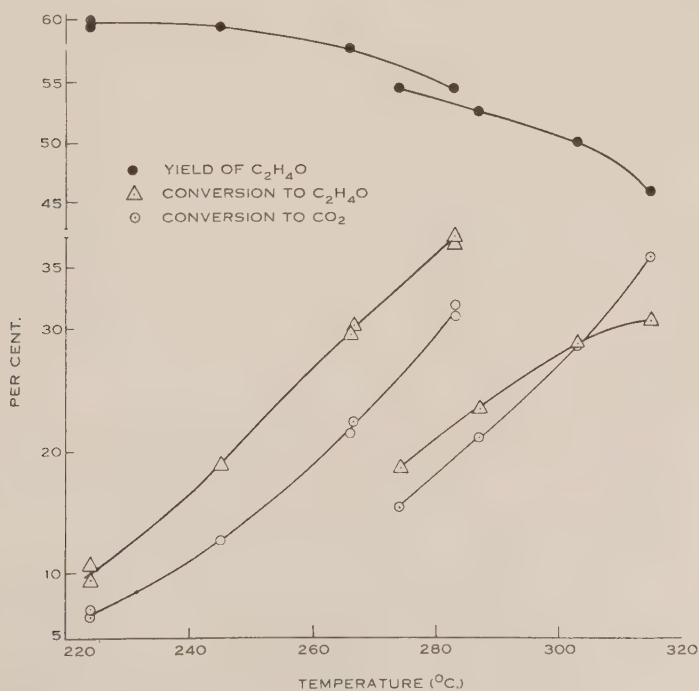


Fig. 6.—Effect of temperature on the reaction.



began to fall and at 300–315 °C. the rate of fall was quite rapid. The conversion to oxide at the higher space velocity approached a maximum at 315 °C.

When the temperature of the reaction was lowered the catalyst retained a higher activity especially towards the production of  $\text{CO}_2$  than would be expected at the lower temperature. After continued use at the lower temperature the catalyst returned to equilibrium and this is shown to some extent in runs 120–122, and by comparing runs 120 and 116. On the other hand, no lag in the change of activity was noticed on increasing the temperature, as is shown by the agreement of the results of the duplicate runs at the lower space velocity.

Apparent activation energies were obtained for the formation of oxide and  $\text{CO}_2$  by plotting for each the conversion in unit contact time against  $1/T$  over the range 225 to 275 °C. The resulting points lay on two straight lines. The apparent activation energy found for the formation of ethylene oxide was 12 kg. cal./g. and for  $\text{CO}_2$  15 kg. cal./g.

A recalculation of the results for the ethylene oxide reaction by the method of Benton(10) gave a value of 11 kg. cal./g. mole.

#### (f) Pressure

Using the second reaction chamber mentioned earlier six runs were made in which the pressure in the reaction chamber was reduced in steps from 762 to 78 mm. The gas composition was 6.25 per cent. ethylene in oxygen, the temperature was kept constant at 260 °C. and the space velocity at the reaction pressure in the limits  $346 \pm 10$ .

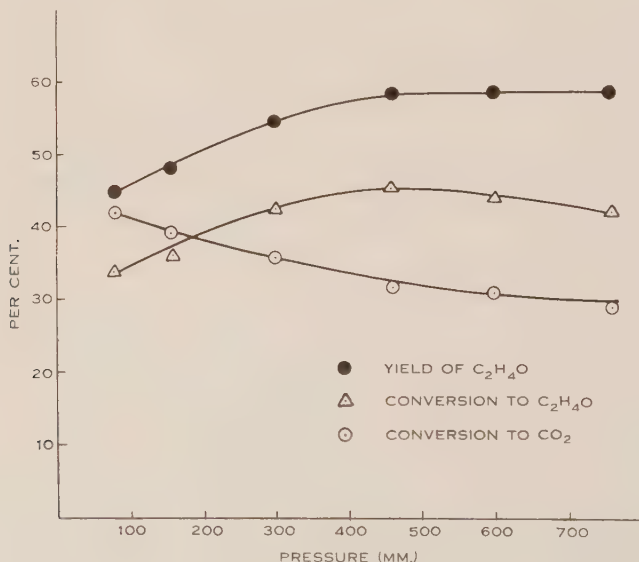


Fig. 7.—Effect of reduced pressure on the reaction.

It is seen from Table 6 and from Figure 7 that as the pressure of the reaction is decreased below 300–400 mm. the proportions of the products change although the total conversion of ethylene is almost constant.

TABLE 6  
EFFECT OF REDUCED PRESSURE

Pressure (mm.)	Conversions		Yield Oxide (%)
	To Oxide (%)	To CO <sub>2</sub> (%)	
762	42.2	28.1	58.5
600	43.8	31.1	58.6
461	45.3	31.8	58.5
302	42.9	35.9	54.4
157	36.4	39.2	48.1
78	34.0	42.1	45.0

(g) *Poisoning of the Catalyst*

Claims made in the patent literature indicate that the addition of small amounts of a substance such as ethylene dichloride to the reacting gases suppresses the oxidation of ethylene to CO<sub>2</sub> and up to a certain concentration increases the conversion to oxide. Above this optimum concentration increasing amounts decrease both conversions to CO<sub>2</sub> and to oxide until the catalyst is totally deactivated. The optimum yield of ethylene oxide observed by Law and Chitwood(11) on adding ethylene dichloride in the example given by them was 60 per cent. of the ethylene oxidized. Although the maximum yields of oxide so far observed in this work are already as high as 60 per cent. it was considered desirable to find how far these yields could be improved on addition of ethylene dichloride vapour.

It was found that the catalyst was easily deactivated by minute amounts of ethylene dichloride, so much so that difficulty was experienced in adding the vapour at a constant measurable rate. Close control was achieved by the constant addition of a carefully controlled amount of a permanent gas saturated with ethylene dichloride. The mixture of gases obtained by the electrolysis of baryta water was used as their rate of production could be easily controlled electrically. By passing them over ethylene dichloride they were saturated with the vapour and admitted to the reaction mixture at a point *U* (Fig. 1). The small amounts of hydrogen introduced with the dichloride would have no appreciable effect on the reaction. In comparative runs at 268–270 °C. with 4.9 per cent. ethylene in the mixture with air, with and without addition of dichloride, it was found that the addition of 0.0003 per cent. by volume of ethylene dichloride vapour decreased both conversions to oxide and to CO<sub>2</sub> from 28.3 and 18.9 to 13.1 and 4.9 per cent. respectively, the yield of oxide thus increasing at the same time from 59.9 to 72.9 per cent.

A series of seven runs was then made with the addition of increasing amounts of ethylene dichloride vapour up to 0.0003 per cent. to find whether smaller

amounts of dichloride caused any increase in the conversion to oxide as observed by previous workers(2, 5). Up to a concentration of 0.0002 per cent. there was little change in conversions or yield of oxide but at this concentration both conversions began to fall and the yield to rise, until at a concentration of 0.0003 per cent. the conversions to oxide and  $\text{CO}_2$  had dropped from initial values of 22.7 and 12.8 to 8.6 and 3.4 per cent. respectively, the yield of oxide rising from 64.0 to 71.8 per cent. This final value is close to the previously observed maximum value, and to a value obtained by Cambron *et al.*(2).

#### (h) Other Constituents of the Reaction Mixture

Several experiments using carbon dioxide in place of the nitrogen of air as a diluent for the reactant gases were compared with runs carried out immediately before and after this substitution. This addition of carbon dioxide decreased the conversion to oxide to two-fifths of the value when using air. Owing to the large amounts of  $\text{CO}_2$  present the amount of  $\text{CO}_2$  produced in the reaction could not be estimated accurately, so that no reliable figures for the yield of oxide can be given.

Cambron *et al.* have observed that small amounts of propylene in the hydrocarbon feed mixture (0.2 per cent. or more) poisoned the catalyst. The effect of the addition of small amounts of propylene was investigated by carrying out comparable runs using (i) pure ethylene, (ii) ethylene containing 0.6 per cent. propylene, and (iii) ethylene containing 5.1 per cent. propylene. The smaller amount of propylene depressed both conversions by approximately 5 per cent. without affecting the yield while 5.1 per cent. propylene caused no further decrease in the conversion to oxide but a small decrease in yield (56 to 54 per cent.) which was apparently due to the oxidation of some of the propylene to  $\text{CO}_2$ . The activity of the catalyst returned to its previous value when subsequent runs with ethylene were made.

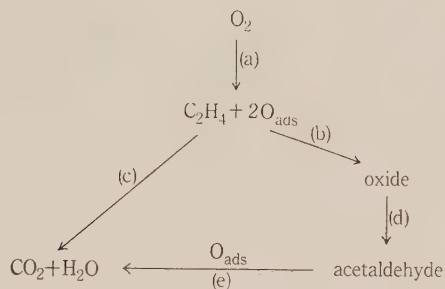
#### (i) Oxidation of Propylene and isoButylene

A number of experiments was made at 250 °C. under similar conditions using reaction mixtures with air containing 5 per cent. ethylene, 5 per cent. propylene, and 5 per cent. *isobutylene* respectively. Whereas ethylene oxide was obtained from ethylene in an 18 per cent. conversion and 58 per cent. yield no oxides could be detected in the oxidation products of propylene and *isobutylene*. The  $\text{CaCl}_2\text{-HCl}$  reagent for oxides and the sensitive test for ethylene oxide using saturated manganous sulphate solution(12) when applied to the condensed liquid products gave negative results. Moreover, determinations of density and refractive index showed that the liquid product in both instances was water, thus eliminating the possibility of any appreciable amount of intermediate oxidation products being present. The relative rates of oxidation based on the total conversion were in the decreasing order ethylene, propylene, *isobutylene*.

### VI. DISCUSSION OF RESULTS

The main stages of the reaction between ethylene, oxygen, and silver giving as products ethylene oxide,  $\text{CO}_2$ , and  $\text{H}_2\text{O}$  have been reported by Twigg(1)

to be (a) the adsorption of oxygen on silver followed by the reaction of ethylene with the adsorbed oxygen to give (b) ethylene oxide or (c)  $\text{CO}_2 + \text{H}_2\text{O}$ . The oxide then undergoes further oxidation to  $\text{CO}_2 + \text{H}_2\text{O}$  by (d) its isomerization on the catalyst to acetaldehyde followed by (e) the oxidation of the acetaldehyde to  $\text{CO}_2 + \text{H}_2\text{O}$ .



The results presented here fit in readily with this scheme and provide some support for it. Thus in keeping with Twigg it has been found that an increase in space velocity hardly changes the ratio of oxide to  $\text{CO}_2$  in the product, which suggests that the main reaction path giving carbon dioxide is not by way of ethylene oxide as the principal intermediate for if this were so one would expect a substantial increase in the ratio oxide/ $\text{CO}_2$  with increase in space velocity.

Twigg has shown that the isomerization of ethylene oxide to acetaldehyde (d) occurs readily on a bare silver catalyst at the reaction temperature. However, in the presence of oxygen the acetaldehyde so formed is rapidly oxidized to  $\text{CO}_2$  and  $\text{H}_2\text{O}$  (e) and never appears in the final reaction products. This second reaction is much the faster and therefore the overall rate of the further oxidation of ethylene oxide is that of its isomerization, the rate of which Twigg considers to be independent of the oxygen concentration on the catalyst. However, it would appear from the above reasoning that the extent of the uncovered surface of the catalyst should influence the isomerization reaction because it determines to what extent the oxide is readsorbed and any variation of it should be apparent by a change in the ratio of oxide to  $\text{CO}_2$ . Thus if the surface covered by oxygen be decreased, then the extent of isomerization reaction should increase, resulting in a decrease in the ratio oxide/ $\text{CO}_2$ . This appears to be a suitable simple explanation for the results obtained when the oxygen content of the reactant gases was varied (Fig. 4) and when the reaction was carried out under reduced pressure (Fig. 6). Both with decrease of oxygen content and with reduction of pressure a decrease in the oxygen-covered surface would be expected and in each case this is accompanied by a substantial decrease in the yield of oxide and in the ratio oxide/ $\text{CO}_2$ .

It has been shown earlier that with a new catalyst the yield of oxide increases considerably until a steady value is reached. As reduction with hydrogen at  $100^\circ\text{C}$ . was the final step in its preparation it may be assumed that initially the catalytic surface was free of oxygen. In contact with the reaction mixture it absorbs oxygen at a high initial rate, the rate diminishing as oxygen is taken



up until an equilibrium is reached. In an adsorption study of oxygen on silver Benton and Drake(13) observed that the time for their system to reach equilibrium was 20 hours at 254 °C. or less than 1 hour at 294 °C. The time taken for the catalyst to reach constant activity at 270 °C. (8–12 hours) roughly corresponds with these values of Benton and Drake and this suggests that after an initial slight drop in overall activity at 240 °C. probably due to sintering, the greater change in activity observed at 270 °C. (Fig. 2) is caused by the establishment of oxygen equilibrium on the catalyst. As this condition is approached and the bare surface of the catalyst decreases, further isomerization of ethylene oxide is diminished with an attendant increase of yield.

The observation that only small changes in yield of ethylene oxide result from changes in space velocity and the ethylene content of the reactant gases is also in accord with the hypothesis that extensive changes in yield were brought about by conditions which altered the amount of adsorbed oxygen on the catalyst. In the space velocity experiments the oxygen content of the reaction mixture was constant and in those concerning ethylene content it was almost constant (between 19.1 and 22.8 per cent.). With a constant partial pressure of oxygen in the gas phase it would be expected that the amount of oxygen adsorbed would also be constant, although there is no means of determining to what extent the increases in reaction rate observed in both series of experiments has affected this amount.

Substantial increases in yield of oxide (from 60 to 73 per cent.) have been obtained by the addition to the reaction mixture of small amounts of ethylene dichloride vapour. This selective poisoning effect to a lesser degree was also observed by saturation of the reaction mixture with water vapour. Langwell(14) has suggested that the action of these selective poisons, ethylene dichloride in particular, is due primarily to the spatial separation of oxygen atoms on the catalyst surface by the adsorption of poison. This lessens the chance of an ethylene molecule encountering adjacent oxygen atoms and being oxidized to  $\text{CO}_2$  according to the mechanism of Twigg. While this hypothesis is an attractive one it is possible that some contribution is made to the suppression of the  $\text{CO}_2$  reaction by the adsorption of the poison on vacant sites on which oxide would otherwise be readsorbed and isomerized.

An interpretation of the results of the effect of temperature on the reaction is not possible owing to the complex changes in the rates of the many processes involved including the rates of desorption of the products. It can be said, however, that rate of isomerization of the oxide will increase with temperature and this could contribute greatly to the reduced yields of oxide obtained with increase of temperature. The apparent activation energies for the two reactions calculated from these results at various temperatures (15 kg. cal. for the reaction giving  $\text{CO}_2$ : 11–12 kg. cal. for the reaction giving oxide) do not agree with the values found by Twigg (29.5  $\pm$  1.5 kg. cal. for reaction giving  $\text{CO}_2$  from static experiments: 27 kg. cal. for the oxide reaction from flow experiments) but agree more closely with his calculated values for ethylene oxide production (16  $\pm$  4 kg. cal.) and direct  $\text{CO}_2$  production (10  $\pm$  3 kg. cal.).

It is surprising that propylene and *isobutylene* have been found to yield no detectable amount of oxides under the same oxidizing conditions as employed for ethylene.\* This, combined with the fact that no other catalyst has to date been shown to oxidize ethylene to its oxide, illustrates the highly specific nature of the reaction between ethylene, oxygen, and silver.

#### VII. ACKNOWLEDGMENTS

Grateful acknowledgment is made to Mr. A. McKenzie for experimental assistance, to Dr. H. H. Hatt and Dr. M. E. Winfield, Division of Industrial Chemistry, C.S.I.R.O., for their interest and cooperation, and to Miss Roberta J. Beckett of the Chemical Physics Section of this Division for taking electron micrographs of several catalysts.

#### VIII. REFERENCES

- (1) TWIGG, G. H.—*Trans. Faraday Soc.* **42**: 284-91 (1946).
- (2) CAMBRON, A., MCKIM, F. L. W., SHANE, G., and MOHUM, W. A.—*Rep. Nat. Res. Coun. Can. Div. Chem.* C 2931-425 (1943).
- (3) *Chemical Age* **51**: 278 (1944).
- (4) REYERSON, L. H., and OPENHEIMER, H.—*J. Phys. Chem.* **48**: 290-5 (1944).
- (5) MCBEE, E. T., HAAS, H. B., and WISEMAN, P. A.—*Industr. Engng. Chem.* **37**: 432-8 (1945).
- (6) MITANI MITSUO and KANO HISAO.—*Mem. Fac. Engng. Hokkaido* **8** (1): 75-106 (1947). (*Chem. Abstr.* **42**: 7246 (1948).)
- (7) TWIGG, G. H.—*Proc. Roy. Soc. A* **188**: 92-141 (1946).
- (8) KERCHOV, F. W.—*Z. Anal. Chem.* **108**: 249-54 (1937).
- (9) LUBATTI, O. F.—*J. Soc. Chem. Ind. Lond.* **52**: 346 T (1932).
- (10) BENTON, A. F.—*Industr. Engng. Chem.* **19**: 494-7 (1927).
- (11) LAW, G. H., and CHITWOOD, K. C.—*Brit. Pat.* 518, 823 (1940).
- (12) LENHER, S.—*J. Amer. Chem. Soc.* **53**: 3739 (1931).
- (13) BENTON, A. F., and DRAKE, L. C.—*J. Amer. Chem. Soc.* **56**: 255-63 (1934).
- (14) LANGWELL, W. H.—*Trans. Faraday Soc.* **42**: 290 (1946).

\* This has been confirmed by later work and has also been found to be so with several higher olefines. An account of this work will be published later.

# THE REACTION BETWEEN ACETONE AND AMMONIA

## II. ISOMERIC OXIMINOKETONES RELATED TO DIACETONAMINE

By N. C. HANCOX\*

[Manuscript received April 20, 1950]

### Summary

The compound 2,2,4,4,6-pentamethyltetrahydropyrimidine hydrate, formed in the reaction between acetone and ammonia at room temperature, is converted by nitrous acid to the 5-oximino derivative in 30 per cent. yield. This oxime is decomposed by acid under mild conditions to give acetone, ammonia, and the salt of 4-amino-4-methyl-2,3-pentanedione 3-oxime. More vigorous acid treatment of the cyclic oxime, or reaction of the open-chain 3-oxime with boiling dilute acid, causes a migration of the oximino group to the 2 position. Evidence is adduced to show that the 2-oxime has the  $\alpha$ - (*anti*) configuration, whereas the 3-oxime has the  $\beta$ - (*syn*) configuration.

### I. INTRODUCTION

In the previous paper of this series, Bradbury, Hancox, and Hatt(1) showed that the reaction between acetone and ammonia at room temperature in the presence of suitable catalysts could be conducted to give 90 per cent. yields of the hydrate of 2,2,4,4,6-pentamethyltetrahydropyrimidine (I). When an attempt was made to prepare an N-nitroso derivative of this compound, there was obtained a low yield of a substance of melting point 168 °C. with the formula  $C_9H_{17}ON_3$ , but the brief study of its properties made at that time indicated clearly that it could not have the constitution expected. The present paper discusses the constitution and reactions of this compound, with particular reference to the substances derived from it by acid degradation.

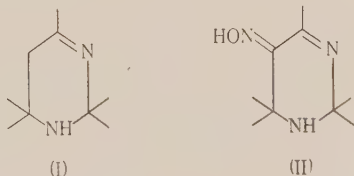
The new compound is readily obtained by treating a mixture of the hydrate of I and excess of sodium nitrite, made up to a slurry with crushed ice, with 1.5 equivalent of very cold 5N hydrochloric acid. The yield of crude material is of the order of 40 per cent., but recrystallization, which is best achieved from methanol, reduces this to 30 per cent., a small amount of unidentified, unstable by-products being removed in the process. The compound separates from methanol with solvent of crystallization, which it rapidly loses in contact with the air; from aqueous solvents a monohydrate can be obtained. The occurrence of hydrates is a common feature among the cyclic bases of this type, which possess the arrangement  $-Me_2C-NH-CMe_2-$ , the pair of gem-dimethyl groups adjacent to the cyclic nitrogen atom apparently being responsible for holding the molecules far enough apart in the crystal lattice to permit the entry of a solvent molecule.

\* Division of Industrial Chemistry, C.S.I.R.O., Melbourne.

Attempts to improve the yield further were unsuccessful; an excess of sodium nitrite appears necessary for best results, but there is no advantage in increasing the amount beyond 1.5 mol. The quantity of acid is rather critical, with the optimum in the range 1.3 to 1.5 mol.; further increase causes a diminution in the yield, which falls to zero if more than 2 mol. of acid is used. The acid must be added as rapidly as possible, with very good cooling. Part of the difficulty in obtaining a high yield arises from the sensitivity of the reduced pyrimidine structure to acids; a characteristic behaviour of this structure is the "drifting end-point" on titration; that is, the compounds titrate initially as monacid bases, but as the titration proceeds decomposition of the cation gradually occurs, leading finally to a stable end-point with 2 mol. of acid. The nitrosation reaction here discussed apparently occurs between nitrous acid and the cyclic cation; the mixture in fact reacts acid to litmus just after the addition of the acid, and remains so during the separation of the product, but becomes alkaline again at the completion of the reaction. Losses inevitably occur by breakdown of both starting material and product during the acid phase of the reaction.

The pure material gives negative Liebermann and diphenylamine reactions; this, and the failure to liberate nitric oxide with concentrated sulphuric acid, make it evident that the compound is not a nitrosamine. Further, the compound has a weakly acidic group, for it dissolves in caustic alkalis, giving deep yellow solutions from which it can be recovered unchanged by saturation with carbon dioxide; this, together with the formation of a monobenzoyl derivative, (I does not benzoylate) suggests the presence of an oximino group.

It has frequently been pointed out that the  $>C=N$  group in heterocyclic compounds may be regarded as an "ammono-carbonyl" group; this group in compound I might therefore be expected to confer reactivity towards nitrous acid on the adjacent methyl or methylene group, and more particularly on the latter (position 5)—by analogy with methyl ethyl ketone. That an oximino group has, in fact, been introduced, and the compound is 2,2,4,4,6-pentamethyl-5-oximino-2,3,4,5-tetrahydropyrimidine (II), is demonstrated by the degradation experiments described below.



Ring fission of II with hydrochloric acid at 30–40 °C. proceeds in the manner typical of tetrahydropyrimidines, leading to equimolar proportions of acetone, ammonium chloride, and the *hydrochloride* of a basic fragment of the composition  $C_6H_{12}O_2N_2$  (III). By addition of II to the calculated quantity of acid at 0 °C., it is possible to isolate a *dihydrochloride* of II, which is decomposed by water at higher temperatures to these same products. If the acid treatment of II is more severe (reflux conditions), the products of degradation are likewise acetone



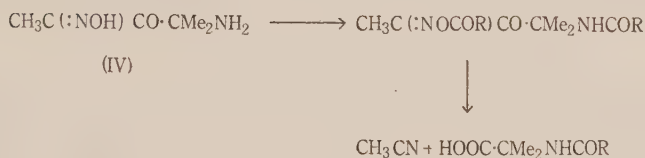
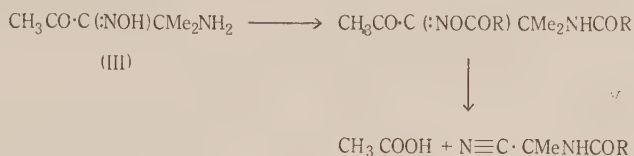
and ammonium chloride, together with the *hydrochloride* of a base (IV), isomeric but not identical with III; and this isomeric hydrochloride is also obtained on brief refluxing of the hydrochloride of III with dilute acid, or on long standing with acid in the cold. Both hydrochlorides separate as crystalline monohydrates.

The isomeric free bases can be liberated from the respective hydrochlorides with sodium bicarbonate; they have closely similar melting points (128–129 °C.), but the mixed melting point shows a large depression. Both are amphoteric, giving colourless solutions in acids, and the deep yellow solutions in alkalis characteristic of 1,2-oximinoketones. Benzoylation gives with each a *dibenzoyl* derivative, that from III melting at 126 °C. and that from IV at 161 °C.; the two differ markedly in solubility. Carbonyl derivatives can also be prepared—a *hydrazone* from each, but the same (*di*)*oxime* from both.

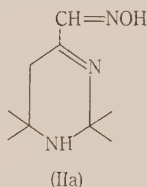
Both III and IV give colour reactions with ferrous salts; IV in weakly alkaline solution gives an intense prussian blue, but III, in striking contrast, gives a deep wine red. With cupric salts the colour reactions of the two compounds are identical—olive green changing to lilac with excess of the organic compound; these colours are stable indefinitely, whereas the iron colours rapidly fade in contact with the air.

The evidence so far adduced is consistent with the postulated structure of the cyclic compound (II), which, by analogy with I, would yield 4-amino-4-methyl-2,3-pentanedione 3-oxime (“*isonitrosodiacetonamine*”) on ring fission. The latter structure admits, in theory, of the existence of two geometrical isomers. However, III and IV are not a pair of such isomers, but are position isomers, which is demonstrated by reactions involving cleavage of their diacyl derivatives at the bond between the carbonyl and the carboxime groups; identification of the fragments provides the proof of the two structures. The *acetyl* derivatives of both III and IV are somewhat unstable and difficult to prepare, because of their tendency to undergo this cleavage. When gently warmed with acetic anhydride, IV undergoes a vigorous exothermic reaction to give a good yield of  $\alpha$ -acetamido*isobutyric* acid; III treated in the same way gives a small amount of  $\alpha$ -acetamido*isobutyronitrile* together with much carbonized material. It is impracticable to isolate the other fragments from this acetylation cleavage, and the corresponding reaction with the benzoyl derivatives gives more useful results; the reaction can be carried out either by gently heating the free oximinoketones with benzoic anhydride until the exothermic reaction is initiated, or by pyrolysing the preformed benzoyl derivatives. Under these conditions, IV cleaves very smoothly to give a good yield of acetonitrile and  $\alpha$ -benzamido*isobutyric* acid; while III gives acetic acid and  $\alpha$ -benzamido*isobutyronitrile*, although the yields are lowered by tar formation and decomposition in other directions (some hydrocyanic acid can be detected among the products). An analogous pyrolytic cleavage has been observed by Blatt and Barnes(2) with the acetate of  $\alpha$ -benzilmonoxime, the products being benzonitrile and benzoic acid (plus acetic acid)—the acetate of the  $\beta$ -oxime resisted cleavage. The same authors obtained pyrolytic cleavage with both forms of benzoin oxime acetate;

the reaction proceeded more smoothly with the  $\alpha$ -form, but the products in each case were benzonitrile and benzoic acid. The reaction is thus a variant of the so-called "second-order Beckmann reaction".



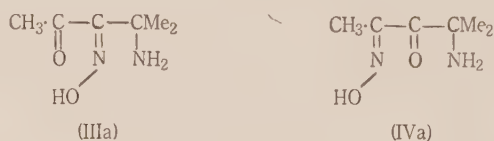
The establishment in this way of the structure of III as 4-amino-4-methyl-2,3-pentanedione 3-oxime also confirms the structure assumed for the cyclic parent (II), and disposes of a possible alternative (IIa).



The reaction responsible for the isomerization of the oximinoketone (III), viz. migration of an oximino group to the adjacent position under the influence of dilute acid, is not without precedent. Borsche(3) obtained from the compound  $p\text{-MeO}\cdot\text{C}_6\text{H}_4\cdot\text{C}(\text{NOH})\text{CO}\cdot\text{CH}_3$  with dilute sulphuric acid a product which he thought to be the result of a Beckmann rearrangement, but which Taylor(4) later showed to be the isomeric oximinoketone  $p\text{-MeO}\cdot\text{C}_6\text{H}_4\cdot\text{CO}\cdot\text{C}(\text{NOH})\text{CH}_3$ ; and a number of other transformations of the type  $\text{Ar}\cdot\text{C}(\text{NOH})\text{CO}\cdot\text{CH}_3 \rightarrow \text{Ar}\cdot\text{CO}\cdot\text{C}(\text{NOH})\text{CH}_3$  have been established by Philipp and Müller(5). The present example appears to be the first observed in the aliphatic series. The reaction has been presumed to take place by hydrolysis and recombination with hydroxylamine under acid conditions; in this connection the resistance of the oximino group of IV to complete removal by acids is noteworthy—prolonged refluxing with concentrated hydrochloric acid causes progressive darkening of the solution, and, although some hydroxylamine can be detected by the benzhydroxamic acid test, no diketone could be isolated.

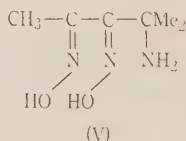
The configurations of the oximino groups in III and IV can be assigned if the conclusions of Barnes and Blatt(6) in regard to the behaviour of acyl derivatives of oximinoketones towards alkali hold in general. These authors show that the acyl derivatives of  $\beta$ - (*syn*) oximinoketones are hydrolysed normally by cold dilute alkali, whereas the derivatives of  $\alpha$ - (*anti*) oximinoketones are

smoothly cleaved (second-order Beckmann reaction) to acid and nitrile. The *ON*-dibenzoyl derivatives of III and IV are little attacked by cold aqueous alkali, but reaction occurs on adding aqueous alkali to their concentrated solutions in alcohol. The derivative of III under these conditions undergoes immediate selective hydrolysis, giving the *N*-monobenzoyl derivative; on the other hand, that of IV is cleaved to  $\alpha$ -benzamidoisobutyric acid. This indicates the  $\beta$ -configuration (IIIa) for III, and the  $\alpha$ -configuration (IVa) for IV.



The colour reaction of IV with ferrous salts gives clear support for the  $\alpha$ -configuration; the formation of complexes with metals is known to be characteristic of *anti*-oximinoketones, and the typical "iron-blue" reaction is commonly regarded as diagnostic of this configuration (cf. 7). On the other hand, the red colour given by III, although non-typical, is to some extent in conflict with the assumption of the  $\beta$ -configuration; such oximinoketones normally are incapable of coordination with metals and so give no colours at all. However, it must be remembered that in III there is an amino group which could become available for coordination when the oximino group is *syn*- to the carbonyl. Exact analogies are lacking, but it is significant that *syn*-phenyl 2-pyridyl ketoxime gives colour reactions with metals(8). The *N*-monobenzoyl derivative of III, with an amido nitrogen adjacent to the oximino group, gives no colour reactions with metals, which seems to indicate that the carbonyl group is not involved in the colour reactions given by III.

The formation of the same dioxime from III and IV is also consistent with the configurations deduced from the evidence given above; entry of a second oximino group could lead in each case to the same *amphi*-dioxime (V). The compound obtained does, in fact, behave as an *amphi*-dioxime towards nickel salts: it gives the pale yellow precipitate characteristic of this configuration instead of the well-known scarlet given by *anti*-dioximes.



If the  $\beta$ -configuration is accepted for III, this compound provides an example of a state of affairs very rare among oximinoketones, viz. the existence of a stable  $\beta$ -form as the only representative of a possible steric pair. Taylor(4), generalizing from the available data on the stereochemistry of oximinoketones, has shown that a compound  $\text{R}_1\text{CO.C}(\text{NOH})\text{R}_2$  exists in two stereoisomeric forms only if  $\text{R}_1$  is an aromatic group and  $\text{R}_2$  is attached to the rest of the molecule

through a tertiary carbon atom. In oximes of this class the  $\alpha$ -configuration is less stable than the  $\beta$ - and can be readily converted to it; but in compounds where  $R_1$  and  $R_2$  do not fulfil the above conditions there is only one form—the  $\alpha$ -form—which is completely stable. He gives as the only known exception ethyl oximinobenzoylacetate, which has the  $\beta$ -configuration. Others may possibly exist among the oximinoketones which fail to give the “iron-blue” reaction, but whose stereochemistry has not been fully investigated. The oximino group in III presumably has its configuration determined by that of the cyclic parent (II), and a plausible reason for the preferential occurrence of the  $\beta$ -form may be found in the steric repulsion of the bulky  $-\text{CMe}_2\text{NH}_2$  group in the adjacent position.

## II. EXPERIMENTAL

Melting points are corrected. Microanalyses were carried out by R. B. Bradbury and L. Oates of these laboratories.

### (i) *5-Oximino-2,2,4,4,6-Pentamethyl-2,3,4,5-Tetrahydropyrimidine (II)*

Freshly prepared 2,2,4,4,6-pentamethyltetrahydropyrimidine hydrate (258 g.; 1.5 mol.) and sodium nitrite (140 g.; 2 mol.) were mixed to a slurry with 150 ml. water and 300 g. crushed ice, and 5N hydrochloric acid (400 ml.; 2 mol.) previously cooled to  $-15^\circ\text{C}$ . was added in one lot with stirring. A transient green colour developed, followed by a vigorous reaction lasting a few seconds accompanied by frothing and the evolution of a small amount of nitrous gases. Some of the product appeared immediately, and the bulk after stirring for several minutes. The mixture was left to stand, with occasional stirring, for 1 hour, then filtered, washed copiously with water, and roughly dried by suction on a Buchner funnel overnight. The yield of crude product was 90–105 g. It was recrystallized, without further drying, from 180 ml. of boiling methanol. It separated as colourless needles which rapidly effloresced and crumbled to powder in the air. Yield 83 g. (30%); m.p.  $166\text{--}168^\circ$  (decomp.).

Found: C, 59.2; H, 9.4; N, 23.2%.

Calculated for  $\text{C}_9\text{H}_{17}\text{ON}_3$ : C, 59.0; H, 9.3; N, 23.0%.

A sample of the freshly prepared crystals obtained from methanol solution gave loss in weight at  $105^\circ\text{C}$ ., 15.2%; calculated for  $\text{C}_9\text{H}_{17}\text{ON}_3 \cdot \text{CH}_3\text{OH}$ : 14.9%.

### (ii) *Hydrate of II*

This separated slowly as granular crystals from aqueous acetone.

Found:  $\text{H}_2\text{O}$ , 9.1% (loss at  $105^\circ\text{C}$ .).

Calculated for  $\text{C}_9\text{H}_{17}\text{ON}_3 \cdot \text{H}_2\text{O}$ : 8.9%.

### (iii) *Dihydrochloride of II*

To ice-cold 5N hydrochloric acid (20 ml.) finely powdered II (9 g.) was slowly added with stirring and cooling. The product, which gradually separated during the addition, was collected, washed with a little ice-water, and dried in a vacuum desiccator. Yield 6.8 g. The compound had no definite melting point, but blackened over a range of  $140\text{--}185^\circ\text{C}$ .

Found: Cl, 23.8%.

Calculated for  $\text{C}_9\text{H}_{17}\text{ON}_3 \cdot 2\text{HCl} \cdot 2\text{H}_2\text{O}$ : Cl, 24.3%.

The compound is freely soluble in water; on gentle warming the solution developed an odour of acetone and deposited crystals of the hydrochloride of III.

### (iv) *Benzoyl Derivative of II*

This was readily obtained by benzoylation under Schotten-Baumann conditions, but purification was difficult on account of its instability; it was best achieved by recrystallizing the crude



material without delay from a small amount of acetone, followed by a second crystallization from benzene/light petroleum. The pure compound forms hexagonal plates, m.p. 105 °C., which decompose on standing for a few days.

Found: C, 66.9; H, 7.6; N, 14.3%.

Calculated for  $C_{16}H_{21}O_2N_3$ : C, 67.0; H, 7.3; N, 14.6%.

(v) *4-Amino-4-Methyl-2,3-Pentanedione 3-Oxime Hydrochloride*

The cyclic oxime (II, 80 g.) was added in small portions to a mechanically stirred solution of 5N hydrochloric acid (180 ml.) so that the temperature was maintained at 35–40 °C. by the heat of reaction. Stirring was continued until crystallization began; this is greatly facilitated when a seed crystal is available from a previous preparation. After cooling in the refrigerator, the solution was filtered, and the crystals washed in turn with small amounts of water and acetone and dried in a vacuum desiccator. Yield of colourless prisms, 78 g. (87%); m.p. 208–210 °C. (decomp.).

The filtrate and aqueous washings smelled strongly of acetone; an aliquot, analysed by Messinger's method, indicated a total of 23 g. (theoretical 24.3 g.). A sample obtained by distillation gave a yellow 2,4-dinitrophenylhydrazone, m.p. 126 °C., undepressed by authentic material. Analysis of the filtrate for ammonia showed 7.3 g. (theoretical 7.45 g.).

For analysis, a sample of the above hydrochloride was recrystallized from water, in which it is moderately soluble.

Found: Cl, 18.0;  $H_2O$ , 9.3%.

Calculated for  $C_6H_{12}O_2N_2 \cdot HCl \cdot H_2O$ : Cl, 17.9;  $H_2O$ , 9.1%.

(vi) *4-Amino-4-Methyl-2,3-Pentanedione 3-Oxime (III)*

The above hydrochloride (75 g.) was intimately mixed with sodium bicarbonate (44 g.) and water (160 ml.) added. The mixture was maintained at approximately 70 °C. in a bath of hot water, with gentle stirring until the evolution of carbon dioxide had ceased (about  $\frac{1}{2}$  hour), and for  $\frac{1}{2}$  hour thereafter. After cooling, the solid was collected by filtration and washed with a little water. The yield of crude dry material was 54 g.; it still contained some inorganic matter. Extraction with boiling acetone and crystallization gave 43 g. (81%) of pure compound as shining colourless prisms, m.p. 128–129 °C.; the melt turned gradually red.

Found: C, 50.0; H, 8.0; N, 19.4%.

Calculated for  $C_6H_{12}O_2N_2$ : C, 50.0; H, 8.3; N, 19.4%.

The compound (III) is readily soluble in cold alcohol, less soluble in cold acetone or ethyl acetate, but very sparingly soluble in ether or chlorinated hydrocarbons. It is moderately soluble in hot water, giving a pale yellow solution; with alkalis, a deep yellow. It is somewhat unstable to heat, undergoing gradual decomposition in air at 100 °C.; solutions in organic solvents darken on prolonged boiling. Treatment with dilute hydrochloric acid in the cold gives a hydrochloride identical with that described above.

(vii) *Colour Reactions of III with Metal Salts*

A drop of dilute ferrous sulphate solution added to a dilute aqueous solution of III produces an intense wine red colour; addition of alkali changes the hue towards purple; acidification destroys the colour. The colour soon fades on exposure to air. A drop of an aqueous solution of III added to dilute copper sulphate solution produces an olive green colour, which changes to lilac as excess of III is added. The colours are stable indefinitely. With nickel salts a pale yellow is obtained.

(viii) *ON-Dibenzoyl Derivative of III*

To a solution of 7 g. of the hydrated hydrochloride of III and 30 g. of anhydrous potassium carbonate, benzoyl chloride was added in small portions with shaking and cooling until the yellow colour disappeared. Shaking was continued until the pasty product solidified. The yield of crude dry material was 12.5 g. (quantitative); two recrystallizations from methanol

gave colourless needles, m.p. 126–126.5 °C. Yield 9 g. (72%). Preparation in like manner from the free base (III) gave similar results. The compound is readily soluble in cold ethanol, less soluble in cold methanol, and insoluble in water.

Found: C, 68.5; H, 5.9; N, 7.9%.

Calculated for  $C_{20}H_{20}O_4N_2$ : C, 68.2; H, 5.7; N, 7.9%.

(ix) *ON-Diacetyl Derivative of III*

One gram of finely powdered III was added cautiously in small portions with good stirring to 1 ml. of pure acetic anhydride. The pasty mixture was then placed in a vacuum desiccator over solid caustic soda and left until the excess anhydride had been removed. The solid remaining was triturated with light petroleum and then recrystallized from ether. Yield 0.85 g. (54%), m.p. 93–94 °C.

Found: N, 12.3; O, 28.0%.

Calculated for  $C_{10}H_{16}O_4N_2$ : N, 12.3; O, 28.1%.

(x) *4-Amino-4-Methyl-2,3-Pentanedione Dioxime (V)*

To a solution of 2 g. of III in 20 ml. of 10% alcoholic sodium hydroxide was added a saturated aqueous solution of 1.2 g. of hydroxylamine hydrochloride. After standing overnight at room temperature, the precipitate of sodium chloride was filtered off and washed with hot alcohol, and the filtrate and washings saturated with carbon dioxide and again filtered. The solution was then evaporated to a small bulk; the product separated as small hexagonal plates. Recrystallization from alcohol gave 1.5 g. (68%); m.p. 170–171 °C.

Found: C, 45.4; H, 8.4; N, 26.3%.

Calculated for  $C_6H_{13}O_2N_3$ : C, 45.3; H, 8.2; N, 26.4%.

(xi) *4-Amino-4-Methyl-2,3-Pentanedione 3-Oxime 2-Hydrazone*

The compound III (7.2 g.) was refluxed for  $\frac{1}{2}$  hour with 10 ml. of 85% aqueous hydrazine hydrate. The initial yellow colour rapidly faded, and on cooling colourless crystals separated. After washing with a little water and drying, the product weighed 7.0 g. (89%), and had m.p. 135–137 °C., unchanged by recrystallization from ethanol. It is sparingly soluble in cold water, more readily in hot; the solution reduces Fehling's solution in the cold.

Found: C, 45.5; H, 9.0; N, 35.3%.

Calculated for  $C_6H_{14}ON_4$ : C, 45.6; H, 8.9; N, 35.4%.

(xii) *4-Amino-4-Methyl-2,3-Pentanedione 2-Oxime Hydrochloride*

The compound II (34 g.) was heated under reflux with 50 ml. of concentrated hydrochloric acid and 50 ml. of water. Crystallization usually occurred spontaneously after 10–20 minutes but occasionally scratching with a glass rod was necessary. The solution after standing overnight in the refrigerator was filtered and the crystals washed with ice-cold water; a further crop, slightly contaminated with ammonium chloride, was obtained by concentrating the mother-liquors. After drying in a vacuum desiccator, the slightly brown, granular crystals weighed 34 g. (92.5%). For analysis, a sample was recrystallized from boiling water, m.p. 228–230 °C. (decomp.). It is less soluble in water than the hydrochloride of III.

Found: Cl, 18.1;  $H_2O$ , 9.0%.

Calculated for  $C_6H_{12}O_2N_2.HCl.H_2O$ : Cl, 17.9;  $H_2O$ , 9.1%.

(xiii) *4-Amino-4-Methyl-2,3-Pentanedione 2-Oxime (IV)*

This preparation was carried out from the hydrochloride in a manner similar to that for III above. Yield of crude, greyish material, 90%; of pure colourless product after crystallization from acetone, 80%, m.p. 129 °C. A mixed m.p. with III showed a depression of 20–30 °C. The compound (IV) separates from acetone as stout square-ended prisms, which, unlike III, are unaffected by heating to 105 °C. in air, and melt with no sign of decomposition. The solubility of IV is somewhat less in all solvents than that of III.

Found: C, 50.0; H, 8.3; N, 19.3%.

Calculated for  $C_6H_{12}O_2N_2$ : C, 50.0; H, 8.3; N, 19.4%.

Prolonged heating of IV with acetone produces a small amount of a yellow crystalline substance, m.p. 115 °C.

Found: C, 58.8; H, 8.9; O, 17.4%.

Calculated for  $C_9H_{16}O_2N_2$ : C, 58.7; H, 8.7; O, 17.4%.

The substance, possibly a Schiff's base, was not further investigated.

(xiv) *Colour Reactions of IV with Metal Salts*

Ferrous sulphate added to a dilute aqueous solution of IV gives an intense prussian blue colour, which is immediately destroyed by acids, and rapidly fades in contact with air. The colour reactions given by copper and nickel salts are similar to those given by III.

(xv) *ON-Dibenzoyl Derivative of IV*

The preparation was similar to that for the dibenzoyl derivative of III, except that 10% sodium hydroxide solution gave better results. The derivative separated and solidified immediately with each addition of benzoyl chloride. After recrystallization from ethanol, a 75% yield of colourless hexagonal tablets, m.p. 161–162 °C., was obtained. The compound is much less soluble in organic solvents than that from III.

Found: C, 68.1; H, 6.0; N, 7.8%.

Calculated for  $C_{20}H_{20}O_4N_2$ : C, 68.2; H, 5.7; N, 7.9%.

(xvi) *ON-Diacetyl Derivative of IV*

This was prepared in the same way as the diacetyl derivative of III; colourless plates from ethyl acetate, m.p. 117–118 °C.

Found: C, 52.7; H, 7.1; N, 12.5%.

Calculated for  $C_{10}H_{16}O_4N_2$ : C, 52.6; H, 7.0; N, 12.3%.

(xvii) *Dioxime (V) from IV*

The preparation was the same as that given above from III except that a longer time (4–5 days) was necessary. The product was identical with that obtained previously, showing no depression in a mixed melting point.

(xviii) *4-Amino-4-Methyl-2,3-Pentanedione 2-Oxime 3-Hydrazone*

This was prepared in the same way as the 3-oxime 2-hydrazone. It formed colourless plates, m.p. 144–145 °C.

Found: C, 45.9; H, 9.0; N, 35.6%.

Calculated for  $C_6H_{14}ON_4$ : C, 45.6; H, 8.9; N, 35.4%.

(xix) *Cleavage of III with Acetic Anhydride*

The compound (III), 1 g., was gently warmed with 1 ml. of acetic anhydride; a violent reaction soon occurred with vigorous boiling of the anhydride, and the mixture darkened. After boiling for a further minute, the mixture was allowed to cool, and kept over sodium hydroxide in a vacuum desiccator to remove excess acetic anhydride. Repeated crystallization of the dark residue from ethyl acetate gave a small yield of almost colourless flakes, m.p. 105–106 °C. This proved to be  $\alpha$ -acetamidoisobutyronitrile (lit. m.p. 105–106 °C.), a mixed melting point with an authentic sample showing no depression.

Found: O, 13.0; N, 21.8%.

Calculated for  $C_6H_{10}ON_2$ : O, 12.7; N, 22.2%.

(xx) *Cleavage of IV with Acetic Anhydride*

This was carried out as described for III above. The mixture did not darken, and from the residue 0.7 g. of pure material, m.p. 196 °C. was obtained after one crystallization from

ethyl acetate. Comparison with an authentic sample of  $\alpha$ -acetamidoisobutyric acid (m.p. 196 °C.) showed the two to be identical.

Found: N, 9.8%; equiv. (monobasic) 149.

Calculated for  $C_6H_{11}O_3N$ : N, 9.7%; equiv. 145.

(xxi) *Pyrolysis of the Benzoyl Derivative of III*

The compound III (4.5 g.) and 7.5 g. of benzoic anhydride, intimately mixed, were placed in a Claisen flask fitted with a thermometer for measuring the internal temperature of the melt, and an adapter leading to a receiver cooled in a dry-ice bath. On application of very gentle heat a reaction set in, and the temperature rose spontaneously to 130 °C.; the mixture darkened considerably. Gentle suction was then applied and heating continued until the temperature reached 160 °C. A distillate of about 1 ml. was obtained. On redistillation, the bulk of this boiled at 116–118 °C., it was identified as acetic acid by conversion to 2-methylbenzimidazole, m.p. 176–177 °C. The crude distillate gave qualitative tests for hydrocyanic acid.

The very dark residue in the flask was dissolved in hot concentrated aqueous ammonia, and allowed to stand overnight. About 1.5 g. of dark, semicrystalline material was filtered off, and purified by repeated crystallization from alcohol, m.p. 168 °C.  $\alpha$ -Benzamidoisobutyronitrile is described in the literature as having m.p. 168–168.5 °C.

Found: C, 70.2; H, 6.4; N, 15.0%.

Calculated for  $C_{11}H_{12}ON_2$ : C, 70.2; H, 6.4; N, 14.9%.

Pyrolysis of the pre-formed dibenzoyl derivative gave like results.

(xxii) *Pyrolysis of the Benzoyl Derivative of IV*

The benzoyl derivative, 10 g., contained in the apparatus described in the previous experiment, was heated in an oil-bath to a few degrees above the melting point. Decomposition then began, and the internal temperature rose to 190 °C. Gentle suction was applied until ebullition of the liquid had ceased. 1.15 g. of distillate was obtained, the bulk of which boiled, on redistillation, at 80–81 °C. It was identified as acetonitrile by reaction with *N*-(hydroxymethyl)-phthalimide to give *N*-(phthalimidomethyl)acetamide; Buc(9), who first described this reaction, gives the m.p. of the latter compound as 213–214.5 °C. (uncorr.); the m.p. here found was 217–218 °C. A sample prepared from authentic acetonitrile had the same melting point, and there was no depression on mixing.

Found: C, 60.7; H, 4.4; N, 12.7%.

Calculated for  $C_{11}H_{10}O_3N_2$ : C, 60.5; H, 4.6; N, 12.8%.

The solid, pale brown residue in the pyrolysis flask was submitted to steam distillation until no more benzoic acid came over; a quantity of brownish crystals remained, which, after filtering and drying, weighed 4.65 g. They were recrystallized from 200 ml. of boiling water (charcoal), and then from acetone, m.p. 202 °C.;  $\alpha$ -benzamidoisobutyric acid is reported(10) to melt at 202 °C.

Found: C, 63.8; H, 6.1; N, 6.8%; equiv. 204.

Calculated for  $C_{11}H_{13}O_3N$ : C, 63.8; H, 6.3; N, 6.8%; equiv. 207.

(xxiii) *N-Monobenzoyl Derivative of III*

The *ON*-dibenzoyl derivative of III (1 g.) was dissolved in 4 ml. of warm alcohol and 10 ml. of cold 10% aqueous sodium hydroxide solution was added. A yellow colour developed immediately, and the small amount of oil which separated soon disappeared. After 5 minutes the solution was saturated with carbon dioxide; the copious precipitate was collected and extracted with hot acetone, and the extract evaporated to dryness. Yield 0.57 g. After recrystallization from methanol it was obtained as colourless prisms, m.p. 172 °C.

Found: C, 63.2; H, 6.2; N, 11.6%.

Calculated for  $C_{13}H_{16}O_3N_2$ : C, 62.9; H, 6.4; N, 11.3%.



The compound dissolved in aqueous alkali giving a yellow solution ; it gave no colour reactions with iron or copper salts. Re-benzoylation gave a product identical with the original dibenzoyl derivative.

(xxiv) *Alkaline Cleavage of the Dibenzoyl Derivative of IV*

The ON-dibenzoyl derivative of IV (1 g.) was dissolved in 15 ml. of boiling alcohol and 10 ml. of cold aqueous 10% sodium hydroxide added. No yellow colour was produced. The solution was cooled and allowed to stand 10 minutes ; then saturated with carbon dioxide. The crystalline material proved to be entirely inorganic (sodium bicarbonate) ; the filtrate was diluted with 10 ml. of water and acidified with hydrochloric acid. From the mixture of solid acids obtained, benzoic acid was removed by sublimation ; the residue (0.4 g.) had m.p. 202 °C. and was identical with  $\alpha$ -benzamidoisobutyric acid obtained in a previous experiment (xxii).

### III. ACKNOWLEDGMENT

The author is indebted to Dr. H. H. Hatt for helpful discussion during the progress of the work.

### IV. REFERENCES

- (1) BRADBURY, R. B., HANCOX, N. C., and HATT, H. H.—*J. Chem. Soc.* **1947** : 1394 (1947).
- (2) BLATT, A. H., and BARNES, R. P.—*J. Amer. Chem. Soc.* **56** : 1151 (1934).
- (3) BORSCHKE, W.—*Ber. dtsh. chem. Ges.* **40** : 742 (1907).
- (4) TAYLOR, T. W. J.—*J. Chem. Soc.* **1931** : 2024 (1931).
- (5) PHILIPP, C., and MÜLLER, S.—*Liebigs Ann.* **528** : 296 (1937).
- (6) BARNES, R. P., and BLATT, A. H.—*J. Amer. Chem. Soc.* **57** : 1330 (1935).
- (7) WHITELEY, M. A.—*J. Chem. Soc.* **83** : 27 (1903) ; TAYLOR, T. W. J.—*J. Chem. Soc.* **1931** : 2024 (1931).
- (8) HUNTRESS, E., and WALTER, H.—*J. Amer. Chem. Soc.* **70** : 3702 (1948).
- (9) BUC, S. R.—*J. Amer. Chem. Soc.* **69** : 254 (1947).
- (10) STEIGER, R. E.—*J. Org. Chem.* **9** : 396 (1944).

## SULPHONAMIDES

### I. "MARFANIL" AND ITS *o*- AND *m*-ISOMERS

By S. J. ANGYAL\* and S. R. JENKIN†

[Manuscript received February 28, 1950]

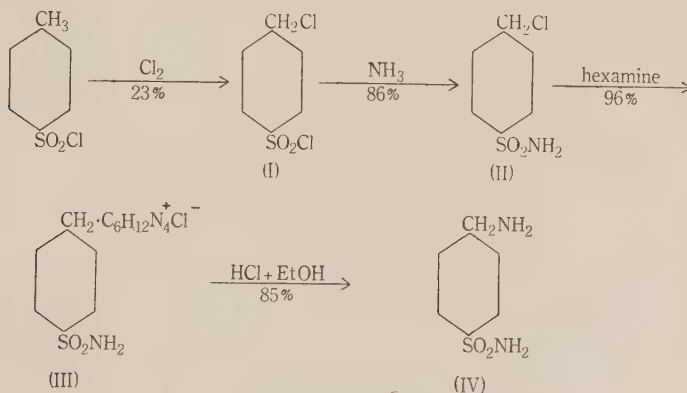
#### Summary

"Marfanil" has been synthesized by a new method from *p*-toluenesulphonyl chloride. *o*- and *m*-Aminomethylbenzenesulphonamide were prepared and found to possess low bacteriostatic activity.

#### I. INTRODUCTION

The drug "marfanil", *p*-aminomethylbenzenesulphonamide (IV) was used extensively by the Germans during World War II for the prevention of gas gangrene infections(1). The interest of the other nations was aroused in 1944 by the announcement(2) that large amounts of marfanil were found amongst medical supplies captured in North Africa. This interest, however, was short-lived because by that time the production of penicillin—a drug more effective against gas gangrene—was well on its way.

In 1944, the manufacture of marfanil in Australia was considered. Since the usual method(3) starts from benzylamine, a substance not freely available in quantity, alternative procedures were tried, and the following series of reactions was successfully carried out:



The chlorination of *p*-toluenesulphonyl chloride has been reported in a patent(4). The method is not very satisfactory and we were unable to improve it, as polychlorination cannot be prevented and to keep it to a minimum the

\* Chemistry Department, University of Sydney.

† Nicholas Pty. Ltd., Melbourne.

reaction must be interrupted before completion. The product can then be separated from unchanged starting material and by-products by distillation at pressures below 1 mm. With the limited equipment available in war-time, this operation, involving a fuming and corrosive mixture, was considered to be impracticable on a large scale. *p*-Toluenesulphonyl chloride could not be chlorinated with sulphuryl chloride in the presence of peroxides. Hence, though the remaining steps of the synthesis proceeded smoothly, this method was abandoned.

The original synthesis(3, 5) of marfanil was then studied. This consists of chlorosulphonation of acetylbenzylamine, treatment of the chlorosulphonyl compound with ammonia, and de-acetylation. The chlorosulphonation was reported(3) to give a 37 per cent. yield which we improved to 47 per cent. In searching for the reason for the low yield we found, by hydrolysing the chlorosulphonation mixture and isolating the insoluble *p*-aminomethylbenzenesulphonic acid, that only 61 per cent. of the acetylbenzylamine was sulphonated in the *p*- position. We looked for a by-product: on treating the crude chlorosulphonation mixture with ammonia and working up the mother liquor of the separated acetylmarfanil, the acetyl derivative of an *isomer*—believed to be *o*—was obtained in 9 per cent. yield. It was converted to the free *amine* and to its *hydrochloride*.

For comparison, the synthesis of the *m*-isomer was undertaken. *m*-Cyano-benzenesulphonamide on catalytic hydrogenation in the presence of ammonia gave only the secondary amine, *m,m'*-bis-sulphamylidibenzylamine, but in the presence of hydrochloric acid, the hydrochloride of the expected *m*-amino-methylbenzenesulphonamide was obtained. This was different from the above by-product which must therefore be the *o*-isomer.

Using a highly virulent strain of *Clostridium welchii* and following the procedure of Lawrence(6), it was found that both the *o*- and *m*-isomers were bacteriostatic at a dilution of 1 in 40. Comparable figures for the *p*-isomer and sodium sulphamerazine were 1 in 16,000 and 1 in 80, respectively.

Because of the important correlation between bacteriostatic power and basic strength, discovered in the acridine series by Albert *et al.*(7), it seemed desirable to determine the ionization constants of the three marfanil isomers. They proved to be identical within experimental error; the values found were: for the *o*-isomer:  $pK_1$ , 8.53;  $pK_2$ , 10.11; for *m*, 8.55 and 10.14; for *p*, 8.52 and 10.08, respectively at 21 °C. In this type of compound, therefore, there is no correlation between basic strength and antibacterial action. The values found for marfanil are not in good agreement, however, with those reported by Goldacre(8), that is 8.18 and 10.23 at 20 °C.

## II. EXPERIMENTAL

All melting points are corrected.

### (a) Preparation of Marfanil from *p*-Toluenesulphonyl Chloride

(i) *p*-Chloromethylbenzenesulphonyl Chloride (I).—A brisk stream of chlorine was passed through commercial *p*-toluenesulphonyl chloride (1675 g.) at 160 °C. for 10 hours. The gain in weight was 32 g. and the chlorine content 28%. On cooling, some *p*-toluenesulphonyl chloride

crystallized out; this was filtered off, and the filtrate distilled *in vacuo*. The distillation was troublesome owing to extensive decomposition. The first fraction, boiling below 134 °C./0.7 mm., consisted mainly of unchanged *p*-toluenesulphonyl chloride; the second fraction was collected up to 154 °C./0.4 mm., leaving a considerable amount of dark residue. This fraction was re-distilled through a short column and yielded 450 g., b.p. 154 °C./0.5 mm., m.p. 50 °C. Yield 23%.

Found: Cl, 31.8%.

Calculated for  $C_7H_6O_2SCl_2$ : Cl, 31.5%.

Two crystallizations from light petroleum gave long colourless needles, m.p. 57.5–58 °C. D.R.—Pat. 234,913(4) reported m.p. 64–65 °C.

(ii) *p*-Chloromethylbenzenesulphonamide (II).—The foregoing compound (25 g.) was added, over a period of 10 minutes, to well-stirred alcoholic ammonia (6.2%; 110 ml.). Ammonium chloride precipitated immediately. After 5 minutes, 200 ml. of water was added, the mixture cooled to 0 °C. and the *product* filtered (19.6 g.), m.p. 161–162 °C. It was recrystallized from alcohol and then melted at 164–165 °C. Yield 86%.

Found: Cl, 17.2%.

Calculated for  $C_7H_8O_2SNCl$ : Cl, 17.25%.

(iii) *p*-Sulphamylbenzylhexaminium Chloride (III).—Quaternary compounds of hexamine are usually prepared in chloroform; but as II is insoluble in this solvent, alcohol was used instead, working in the cold because heating in this solvent caused decomposition of the quaternary salt. To a solution of II (5 g.) in alcohol (75 ml.), hexamine (3.75 g.) in water (3 ml.) was added. After 9 days, 8.1 g. of flaky *crystals* were collected; they decomposed about 180 °C. and apparently retained solvent tenaciously because the chlorine content was always low. Yield 96%.

Found: Cl, 9.4–9.7%.

Calculated for  $C_{13}H_{20}O_2SN_5Cl$ : Cl, 10.3%.

(iv) *Marfanil* (IV).—The quaternary salt (5.2 g.) was heated under reflux with alcohol (40 ml.) and concentrated hydrochloric acid (10 ml.) for 30 minutes, then evaporated to dryness on the water-bath. The residue was dissolved in the minimum amount of water, and 20% sodium hydroxide solution was added with heating and stirring until the pH was 9.4; after cooling to 0 °C., marfanil (2.2 g.) crystallized slowly, m.p. 140–142 °C. Sodium hydroxide (c. 3 g.) was dissolved in the mother liquor when, on cooling, the sodium salt of marfanil separated. This salt is only very slightly soluble in strong solutions of sodium hydroxide. The salt was filtered, dissolved in the minimum amount of water, and the solution brought to pH 9.3 with hydrochloric acid, whereupon an additional amount of marfanil (0.1 g.) separated. Crystallization of the combined products from water (10 ml.) gave 1.6 g. marfanil, m.p. 148 °C. Yield 85%.

#### (b) Preparation of Marfanil from *N*-Benzylacetamide

(i) *Acetylmarfanil*.—*N*-Benzylacetamide (977 g.) was melted and added dropwise to well-stirred chlorosulphonic acid (2100 ml.), cooled to below 20 °C. The mixture was heated to 70 °C., held at this temperature for an hour and then cooled to 0 °C. On the next day, the mixture was poured, with vigorous stirring, on to 10 kg. of ice. A viscous oil separated. After half an hour the aqueous layer was decanted, and the oil added slowly to a well-stirred mixture of concentrated ammonia (2 l.) and water (1 l.), keeping the temperature below 40 °C. Crystallization started when about half of the oil had been added. The mixture was warmed to 75 °C. with stirring and then kept at 0 °C. overnight. Filtration and recrystallization from water gave 704 g. of acetylmarfanil, m.p. 170–171 °C. Yield 47%.

(ii) *The Degree of p*-Sulphonation.—A chlorosulphonation mixture, prepared from 2.8 g. of *N*-benzylacetamide as given above, was poured into water (70 ml.) and refluxed for 5 hours. On cooling, *p*-aminomethylbenzenesulphonic acid(9) crystallized in glittering flakes (2.15 g.).



This yield was considered to be indicative of *p*-sulphonation because the solubility of this acid in cold water is low (less than 1 in 300). The amino acid is neutral but can be titrated with alkali in the presence of formaldehyde. Yield 61%.

Found: Equiv. wt. 185.

Calculated for  $C_7H_9O_3SN$ : Equiv. wt. 187.

The same compound was also obtained from marfanil in 75% yield on refluxing with 70% sulphuric acid for 2 hours.

(iii) *Marfanil Sulphate*.—To marfanil base in alcohol the equivalent amount of concentrated sulphuric acid was added. The *sulphate* precipitated immediately. It was recrystallized from 7 parts of 70% alcohol. The melting point varied with the rate of heating; in a capillary heated at 2 °C. per minute, it melted at 254–255 °C. with decomposition. Salts of marfanil can be titrated with alkali in the presence of formaldehyde, using thymol blue as an indicator. The equivalent weight of the sulphate was determined by this method.

Found: Equiv. wt. 236.

Calculated for  $C_{14}H_{20}O_4S_2N_4 \cdot H_2SO_4$ : Equiv. wt. 235.

#### (c) Preparation of the *o*-Isomer

(i) *o*-Acetylaminomethylbenzenesulphonamide.—To obtain seed crystals, some acetylmarfanil mother liquor (Section II (b) (i)) was brought to dryness *in vacuo*, extracted with alcohol which left ammonium chloride undissolved, and the extract concentrated and kept in the refrigerator with frequent scratchings for weeks. The solid thus obtained was used to inoculate 2 l. of the mother liquors. After 3 weeks at 0 °C., 65 g. of crude product separated. Two crystallizations from water gave the pure compound, m.p. 137–138 °C. Yield 9%.

Found: N, 12.3%.

Calculated for  $C_9H_{12}O_3SN_2$ : N, 12.3%.

(ii) *o*-Aminomethylbenzenesulphonamide.—The above compound (20 g.) was refluxed for 8 hours with alcohol (200 ml.) and concentrated hydrochloric acid (200 ml.), then evaporated to dryness on the water-bath. The residue was dissolved in 50% alcohol (30 ml.); on addition of absolute alcohol (120 ml.) the *hydrochloride* of *o*-aminomethylbenzenesulphonamide crystallized, m.p. 238–239 °C. Yield 16 g. (82%).

Found: N, 12.3%.

Calculated for  $C_7H_{11}O_2SN_2Cl$ : N, 12.6%.

To prepare the free base, the hydrochloride (4.0 g.) was dissolved in water, anhydrous sodium carbonate (1.0 g.) was added, and the mixture evaporated to dryness. The residue was extracted with absolute alcohol, the solvent removed, and the residual oil dissolved in water (5 ml.); on cooling, the *base* crystallized (1.7 g.), m.p. 83–84 °C. It is very soluble in water and in most of the common organic solvents. On an attempt to crystallize it from acetone (10 parts), the *isopropylidene compound* separated, m.p. 136 °C. Yield 83%.

Found: N, 12.3%.

Calculated for  $C_{10}H_{14}O_2SN_2$ : N, 12.4%.

#### (d) Preparation of the *m*-Isomer

(i) *m*-Cyanobenzenesulphonamide.—Metanilamide (17.2 g.) was dissolved in a hot mixture of concentrated hydrochloric acid (23 ml.) and water (50 ml.). The solution was rapidly cooled to 0 °C., whereupon the hydrochloride of the base crystallized, and diazotized with a solution of sodium nitrite (7.6 g.) in water (20 ml.). The mixture was then added, over a period of 20 minutes, to a previously prepared solution of cuprous chloride (12 g.) and sodium cyanide (16 g.) in water (100 ml.). The temperature during this addition was kept below 20 °C., then slowly raised to 70 °C. After cooling, the precipitate was filtered, dried, and extracted twice with a boiling mixture of benzene (180 ml.) and alcohol (20 ml.). The combined extracts were evaporated to dryness and the residue crystallized from water (100 ml.), giving 9.0 g. (50%) of *m*-cyano-benzenesulphonamide, m.p. 148–150 °C. Wallach and Huth(10), who prepared this compound from *m*-sulphobenzoic acid, reported m.p. 151–152 °C.

(ii) *m,m'*-Bis-sulphamyldibenzylamine.—The above compound (5.0 g.) was hydrogenated in the presence of methanol (120 ml.), concentrated ammonia (8 ml.), and 10% palladium charcoal (0.6 g.) at atmospheric temperature and pressure. Two mols. of hydrogen was absorbed in 24 hours. The solution was filtered, evaporated to dryness, and the residue crystallized twice from water, giving fine needles of the compound, m.p. 176.5–177 °C. Yield 3.0 g. (60%).

Found: N, 11.7%.

Calculated for  $C_{14}H_{17}O_4S_2N_3$ : N, 11.8%.

The hydrochloride was crystallized from 2 parts of water by the addition of 10 parts of alcohol, m.p. 216–218 °C.

Found: N, 10.8%.

Calculated for  $C_{14}H_{18}O_4S_2N_3Cl$ : N, 10.7%.

(iii) *m*-Aminomethylbenzenesulphonamide.—*m*-Cyanobenzenesulphonamide (2.0 g.), alcohol (35 ml.), water (30 ml.), concentrated hydrochloric acid (2.0 ml.), and 10% palladium charcoal (0.5 g.) were hydrogenated at atmospheric temperature and pressure. Two mols. of hydrogen was taken up in 15 hours. The solution was filtered, evaporated to dryness, and the residue triturated with alcohol (10 ml.): 1.4 g. (57%) of the crude hydrochloride was obtained. Crystallization from water (1.5 ml.) with the addition of alcohol (9 ml.) gave the pure hydrochloride, m.p. 232 °C.

Found: N, 12.3%.

Calculated for  $C_7H_{11}O_2SN_2Cl$ : N, 12.6%.

The free base, prepared in the same way as described for the *o*-isomer, was a viscous oil which did not crystallize.

#### (e) *pK* Determinations

The ionization constants were determined by potentiometric titrations: 0.005 M solutions of the hydrochlorides (200 ml.) were titrated with 0.1 N potassium hydroxide solution. The *pK* values were calculated from three pH readings and were corrected for hydrolysis of the salt formed. Each determination was carried out three times.

### III. ACKNOWLEDGMENTS

Our gratitude is due to the Directors of Nicholas Pty. Ltd. for their permission to publish this paper; to Miss B. Naylor, microanalyst, University of Sydney, for some of the analyses; to Mr. A. Bryson, Sydney Technical College, and Mr. C. L. Angyal, University of Sydney, for their assistance in the *pK* determinations; and to Mr. C. C. Kuchel, Nicholas Pty. Ltd., for carrying out the bacteriological tests.

### IV. REFERENCES

- (1) KLARER, J.—*Klin. Wschr.* **20**: 1250 (1941).
- (2) ANON.—*Lancet* **1**: 635 (1944).
- (3) BERGEIM, F. H., and BRAKER, W.—*J. Amer. Chem. Soc.* **66**: 1459 (1944).
- (4) BADISCHE ANILIN- und SODAFABRIK. D.R.-Pat. 234,913; *Friedländer* **10**: 117 (1910-12).
- (5) MILLER, E., SPRAGUE, J. M., KISSINGER, L. W., and MCBURNEY, L. F.—*J. Amer. Chem. Soc.* **62**: 2101 (1940).
- (6) LAWRENCE, C. A.—*J. Bact.* **49**: 149 (1945).
- (7) ALBERT, A., RUBBO, S. D., GOLDAIRE, R. J., DAVEY, M. E., and STONE, J. D.—*Brit. J. Exp. Path.* **26**: 160 (1945).
- (8) GOLDAIRE, R. J.—*Nature* **154**: 796 (1944).
- (9) JENSEN, K. A., SCHMITH, K., BRANDT, P., LAURITSEN, M., and HANSEN, O. R.—*Z. physiol. chem.* **280**: 37 (1944).
- (10) WALLACH, O., and HUTH, T.—*Ber. deutsch. chem. Ges.* **9**: 427 (1876).

# DUNNIONE AND RELATED NAPHTHOQUINONES

## I. THE *isodunnione* SERIES

By R. G. COOKE\* and T. C. SOMERS\*

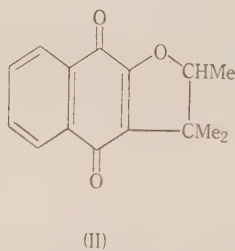
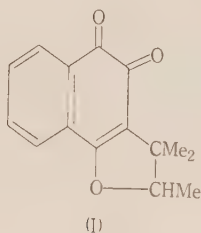
[Manuscript received March 2, 1950]

### Summary

Synthesis of hydroxyhydroisodunnioniol proves the structure of this compound, and of  $\alpha$ - and  $\beta$ -isodunnione. Experiments on the synthesis of isodunnioniol do not support the structure previously assigned to it, and the alternative formula, 2-hydroxy-3-(1,2-dimethylallyl)-1,4-naphthoquinone, is suggested. Evidence is presented for the occurrence of steric inhibition of resonance in certain 2-hydroxy-3-alkylene-1,4-naphthoquinones. Synthetic homologues of the lapachol and dunnione series are described.

## I. INTRODUCTION

Dunnione, the orange-red pigment of *Streptocarpus dunnii* Mast., first investigated by Price and Robinson(1-3), is remarkable for the ease with which it is converted into a number of isomers, and other derivatives, by mild treatment with acid or alkali. In this it resembles lapachol (VIII), another natural colouring matter, with which it is isomeric. Dunnione is evidently a  $\beta$ -naphthoquinone derivative, and Price and Robinson suggested the structure I as being most consistent with the results of mild oxidation reactions. Of the known isomers of the pigment,  $\alpha$ -dunnione is most closely related, and since the two compounds are readily interconvertible the structure II was suggested for  $\alpha$ -dunnione.



The ease with which other isomers may be obtained from I and II lessens the significance of the oxidation reactions on which the structures are based. It is therefore desirable to obtain unequivocal proof by synthesis of all the compounds concerned. This paper is the first of a series dealing with the synthesis of dunnione and related compounds.

To avoid the complications associated with the optical activity of dunnione and some of its transformation products, attention is first directed to the inactive

\* Department of Chemistry, University of Melbourne.

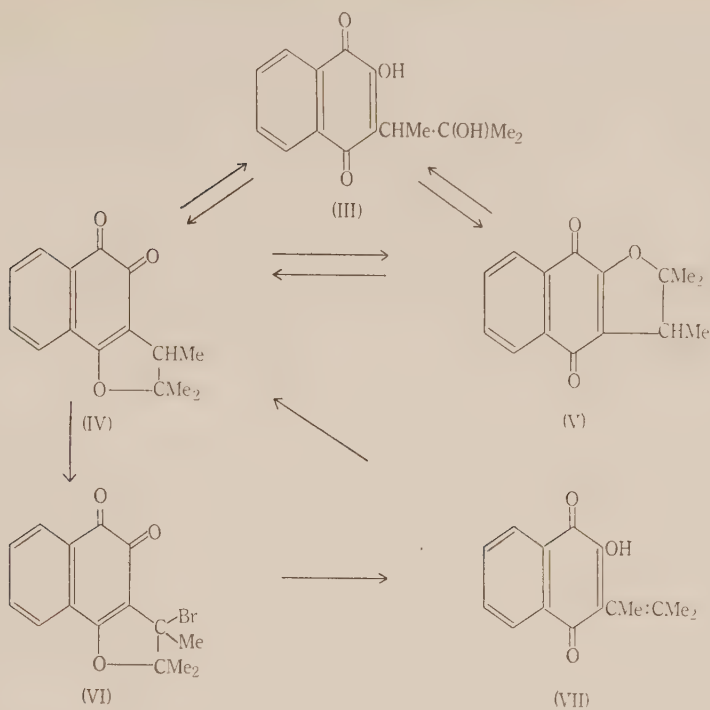


Fig. 1

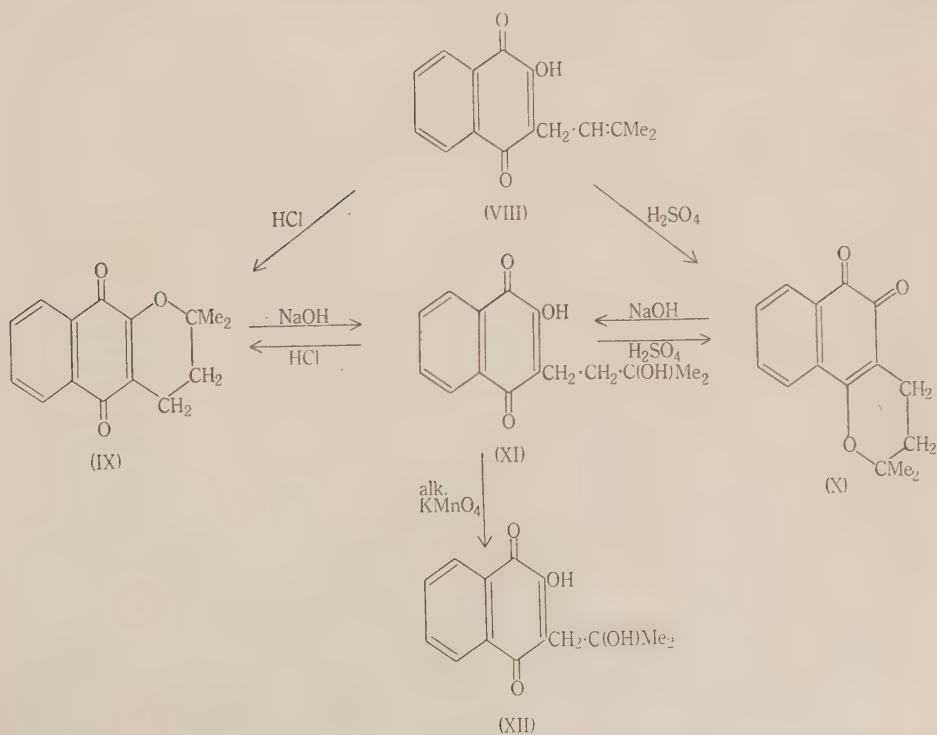


Fig. 2



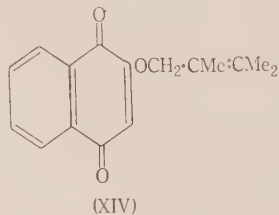
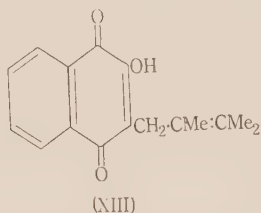
*isodunnione* series. The relationship between these compounds and the structures tentatively assigned to them by Price and Robinson(2, 3) are shown in Figure 1.

Treatment of  $\alpha$ -dunnione (II) with hot concentrated sulphuric acid gives  $\beta$ -*isodunnione* (IV) readily interconvertible with  $\alpha$ -*isodunnione* (V). Treatment of either IV or V with sodium hydroxide solution results in opening of the five-membered ring, and hydroxyhydroisodunniol (III) may be isolated by careful neutralization. Strong acids cyclize III to either IV or V depending on the conditions. Bromination of IV gives  $\gamma$ -bromo- $\beta$ -*isodunnione* (VI), which on treatment with zinc dust and alkali is debrominated. After filtration and aeration of the alkaline mixture, the addition of mineral acid precipitates *isodunniol* (VII). Treatment of VII with concentrated sulphuric acid reforms IV.

Some aspects of the chemistry of lapachol (VIII) are also important in the study of dunnione and its derivatives. Thus Hooker has shown(4-6) that VIII may be converted to IX, X, XI, and XII, as shown in Figure 2.

(a) *Synthesis of Hydroxyhydroisodunniol (III)*

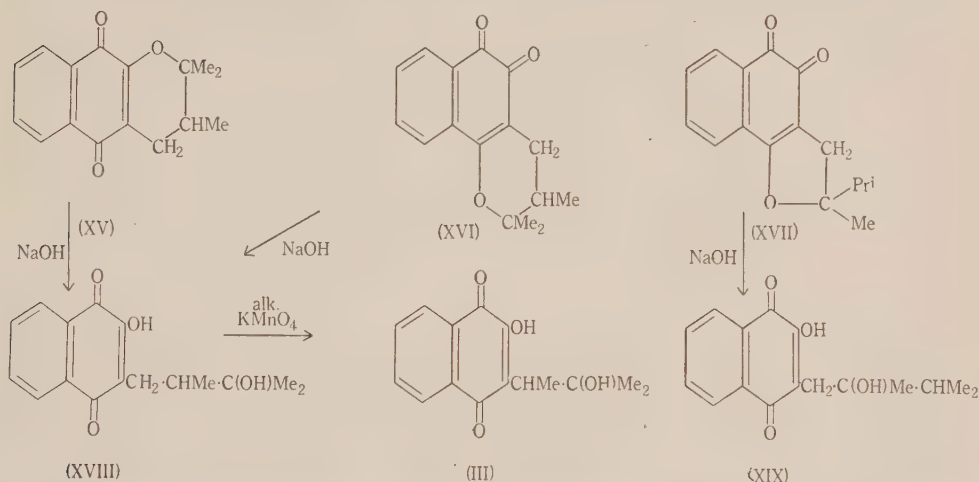
Reactions similar to those of lapachol are given by the homologous compound, 2-hydroxy-3-(2,3,3-trimethylallyl)-1,4-naphthoquinone (XIII), which is readily obtained, together with 2-(2,3,3-trimethylallyloxy)-1,4-naphthoquinone (XIV) by condensation of trimethylallyl bromide with the silver salt of 2-hydroxy-1,4-naphthoquinone. This is an application of the general method of Fieser(7, 8) for the preparation of allylhydroxynaphthoquinones.



Markownikoff's rule predicts the formation of six-membered rings in the cyclization of VIII, but in the case of XIII the rule is not so clearly discriminating. The concept of hyperconjugation and the cyclization of various allyl phenols studied in the course of work on vitamin E(9) suggest the probable formation of mixtures of five and six-membered ring compounds from XIII, with the pyrano derivatives predominating. Treatment of XIII with hot concentrated hydrochloric acid in glacial acetic acid solution gives only one compound, which is evidently  $\beta\gamma\gamma$ -trimethyldihydropyrano-1,4-naphthoquinone (XV).<sup>\*</sup> Cold concentrated sulphuric acid, however, cyclizes XIII to a mixture of two compounds,  $\beta\gamma\gamma$ -trimethyldihydropyrano-1,2-naphthoquinone (XVI) and  $\beta$ -isopropyl- $\beta$ -methyldihydrofurano-1,2-naphthoquinone (XVII). The proportions in which these two compounds are formed depend to some extent on the time of contact with

<sup>\*</sup> The nomenclature for compounds of this type was suggested by Fieser; see Hooker(6).

the concentrated sulphuric acid, but XVI always appears to preponderate. When the reaction period is sufficiently long, only XVI can be isolated. This behaviour could be explained by assuming that initially a hydrogen ion can be added with comparable facility to either of the unsaturated carbon atoms, and that eventually the less stable carbonium ion must isomerize, either by a "hydride shift"(10), or by the establishment of an equilibrium, so that the more stable six-membered ring is obtained. Alternatively a mesomeric carbonium ion might be involved. Treatment of XV and XVI with sodium hydroxide leads to opening of the ring and formation of 2-hydroxy-3-(2,3-dimethyl-3-hydroxybutyl)-1,4-naphthoquinone (XVIII). Similarly XVII can be converted into 2-hydroxy-3-(2,3-dimethyl-2-hydroxybutyl)-1,4-naphthoquinone (XIX).



These structures for the dihydroxyquinones are based on the results of oxidation with alkaline hydrogen peroxide, XVIII and XIX yielding acetone and methyl isopropyl ketone respectively.

Hooker oxidation(6) of XVIII results in a change analogous to XI→XII, and the product is hydroxyhydroisodunnione (III), the identity of the product being shown by mixed melting point with the specimen originally obtained from natural dunnione by Price and Robinson(3). Since III can be obtained by the action of alkali on  $\beta$ -isodunnione (IV) or  $\alpha$ -isodunnione (V), its synthesis confirms the structures assigned by Price and Robinson to these three compounds. The synthetic III also gives IV and V by appropriate treatment with acids.

The formation of  $\beta$ -isodunnione from  $\alpha$ -dunnione is most logically interpreted as a retopinacolinic rearrangement (Fig. 3). Consequently the establishment of the formula for  $\beta$ -isodunnione (IV) supports the structures I and II put forward for dunnione and  $\alpha$ -dunnione.

#### (b) Tetramethyldihydrofuranonaphthoquinones

The structures of III and related compounds which were discussed in the preceding section are substantiated by a study of products obtainable from the allyloxynaphthoquinone (XIV). Fieser has shown(7, 8) that such compounds

undergo Claisen rearrangement relatively easily. XIV is outstanding in this respect, a short period of boiling in ethanol giving a practically quantitative rearrangement to *2-hydroxy-3-(1,1,2-trimethylallyl)-1,4-naphthoquinone* (XX). Although Claisen rearrangement of  $\gamma$ -substituted allyl ethers may occur with attachment at either the  $\beta$  or the  $\gamma$  carbon atom(11), XX will be formed from

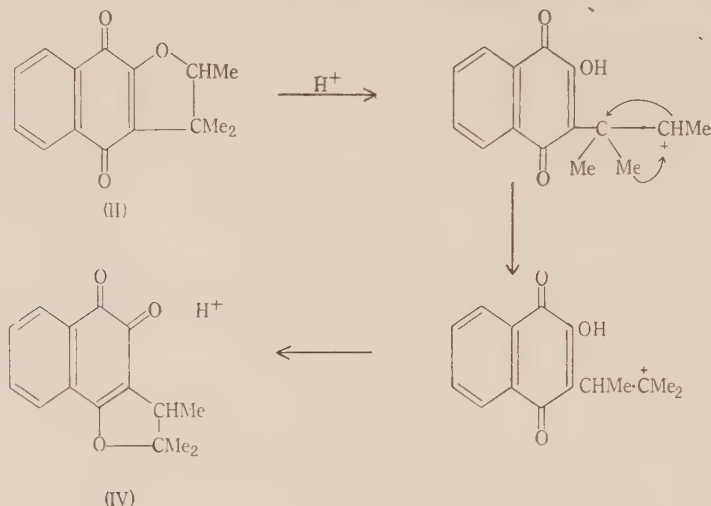
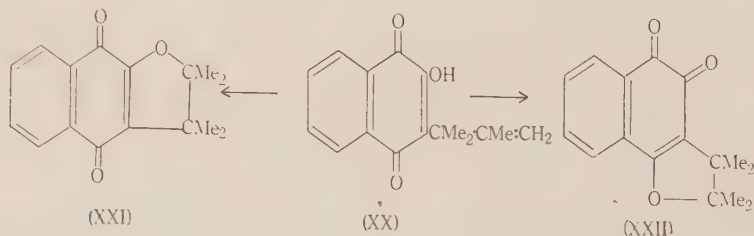


Fig. 3

XIV whichever course the reaction takes. Suitable treatment with acids converts XX into  $\alpha\alpha\beta\beta$ -tetramethyldihydrofurano-1,4-naphthoquinone (XXI) and  $\alpha\alpha\beta\beta$ -tetramethyldihydrofurano-1,2-naphthoquinone (XXII). These compounds are closely related to the dunniones and isodunniones, but are more resistant to the action of acid and alkali. Their structures follow from Markownikoff's rule and from the formation of acetone by oxidation of XXII with alkaline hydrogen peroxide.



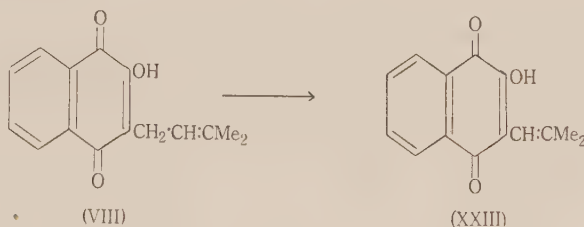
No retropinacolinic rearrangement of XXI would be expected, and it is found to give only XXII when heated with concentrated sulphuric acid under conditions which result in conversion of  $\alpha$ -dunnione (II) to  $\beta$ -isodunnione (IV).

The stability of the heterocyclic system in XXI is shown by the observation that, when XXII is boiled in aqueous alkali, XXI separates from the solution. Under these conditions both dunnione (I) and  $\alpha$ -dunnione (II) dissolve rapidly with rearrangement to another isomer called *allodunnione*. However, cold

dilute alkali converts dunnione to  $\alpha$ -dunnione, together with some *allodunnione*. According to Price and Robinson(2), the formation of *allodunnione* involves fission of the quinone ring. The reaction is now known to occur with various hydroxynaphthoquinones(12, 13) and it will be discussed fully in a later paper of this series.

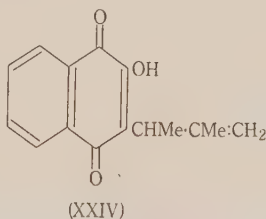
(c) *The Structure of isoDunnioI*

Hooker has shown(6, 14) that 2-hydroxy-3-alkylene-1,4-naphthoquinones having the side-chain double bond conjugated with the quinone ring are usually red, or orange-red, and give violet solutions in alkali. Compounds having saturated side-chains, or with a double bond in some other position, are yellow, and their alkaline solutions are red. Thus Hooker oxidation(6) of lapachol (VIII), which is yellow, gives the orange-red dimethylvinylhydroxynaphthoquinone (XXIII).



Measurements of the absorption spectra of these and related compounds by Cooke, Macbeth, and Winzor(15) clearly distinguish between the two types of structure.

The method of preparation of *isodunnioI* suggests the structure VII, but the yellow colour of this compound and the red solutions it gives with alkali seem more consistent with the alternative allyl structure XXIV.



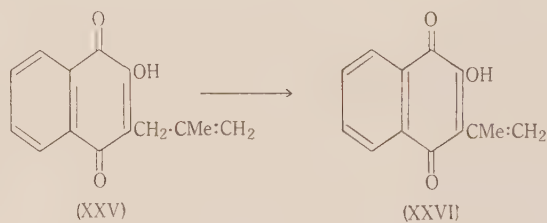
These colour differences, however, provide insufficient grounds for rejection of structure VII; there is another factor which may modify the colour of such compounds. Models indicate that significant steric inhibition of resonance may be anticipated in molecules of this type. The proximity of the  $\alpha$ -methyl group of the side-chain to the two oxygen atoms on the quinone ring should prevent a fully coplanar structure of the conjugated system, and a lighter colour would be expected. Such an effect is evident in the absorption spectra of certain  $\alpha$ -methylstyrenes(16, 17). A decision between structures VII and XXIV for *isodunnioI* can best be obtained by synthesis.



The most obvious method for obtaining VII is by application of the Hooker oxidation to XII, giving a change analogous to VIII→XXIII, but the original one-stage process using alkaline permanganate(6) is not successful. The two-stage modification of Fieser and Fieser(18) evidently produces the required compound, but it is difficult to crystallize, is unstable, and has not been obtained pure. The crude material is readily cyclized to  $\beta$ -isodunnione (IV) by treatment with concentrated sulphuric acid. The light colour (golden-yellow) of the partly purified material indicates that the conjugation is inhibited by the steric effect already discussed. However, other properties of this synthetic product, so far as they are known, suggest that it is not identical with isodunniol, and so support the alternative structure XXIV for that compound.

An investigation of the absorption spectra of isodunniol and authentic representatives of the two structural types might be expected to yield some evidence of value. Since the synthetic VII has not been obtained pure, another representative of this group of compounds has been prepared for this purpose.

By alkylation of 2-hydroxy-1,4-naphthoquinone with methylallyl iodide, 2-hydroxy-3-(2-methylallyl)-1,4-naphthoquinone (XXV) is obtained. Oxidation of XXV by the two-stage process of Fieser and Fieser(18) gives 2-hydroxy-3-isopropenyl-1,4-naphthoquinone (XXVI), which is the simplest of the sterically hindered naphthoquinones.



The compound XXVI is more stable than VII and is readily purified. Its colour (orange-yellow) is consistent with steric inhibition of conjugation, although the effect is probably less in this case than it is with the more bulky side-chain of VII. It is evident from the data in Table 1, however, that the absorption spectrum of isodunniol does not allow a decision to be made between the structures VII and XXIV. There is close resemblance with the spectra of both lapachol (VIII) and the isopropenylhydroxynaphthoquinone (XXVI). The influence of the steric effect on the absorption of XXVI can be appreciated by comparison with the spectrum of the dimethylvinylhydroxynaphthoquinone (XXIII). Another example of this effect is described by Cram(25).

The structure of XXVI is confirmed by hydrogenation to the known 2-hydroxy-3-isopropyl-1,4-naphthoquinone, and by its behaviour with strong acids which hydrolyse it to acetone and 2-hydroxy-1,4-naphthoquinone. Terminal addition of a hydrogen ion would be expected (Markownikoff rule), and since cyclization to a four-membered ring would be difficult the reaction is apparently completed by attachment of a hydroxyl ion and fission of what is, in effect, a  $\beta$ -ketol type addition product of two ketones (XXVII). The process is illustrated in Figure 4.

TABLE I  
ABSORPTION SPECTRA (IN ETHANOL)

Compound	$\lambda$ (m $\mu$ )	log $\epsilon$	$\lambda$ (m $\mu$ )	log $\epsilon$	$\lambda$ (m $\mu$ )	log $\epsilon$	$\lambda$ (m $\mu$ )	log $\epsilon$
<i>iso</i> Dunniol .. ..	251	4.31	282	4.13	331	3.43	380*	3.11
<i>iso</i> Propenylhydroxynaphthoquinone .. ..	250	4.25	274	4.14	333	3.44	380	3.06
Lapachol† .. ..	252	4.38	278	4.28	331	3.43	382	3.17
Dimethylvinylhydroxynaphthoquinone† ..	—	—	265	4.39	317	3.50	421	3.32

\* Figures in italics represent inflexions.

† Cooke, Macbeth, and Winzor(15).

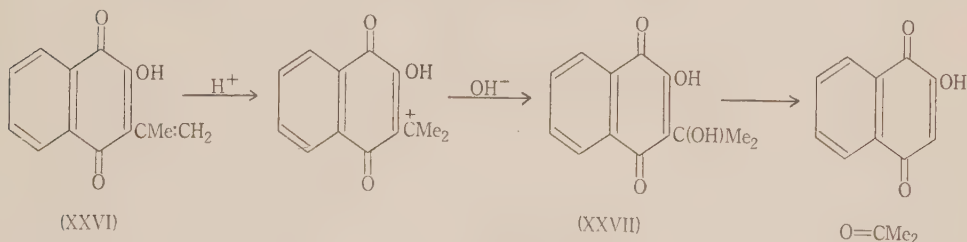
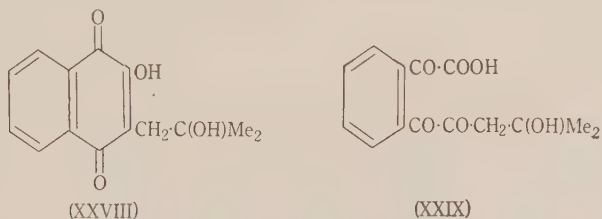


Fig. 4

It is significant that 2-hydroxy-3-(2-hydroxyisobutyl)-1,4-naphthoquinone (XXVIII) gives abnormal results when oxidized with alkaline permanganate. Instead of the expected degradation to XXVII, much of the starting material was recovered unchanged, and "there were some indications of the formation of 2-hydroxy-1,4-naphthoquinone"(19).



It seems likely that the strongly alkaline medium of the Hooker oxidation could also catalyze the splitting of acetone from the  $\beta$ -ketol (XXVII) but this step might well occur with the open-chain ketol (XXIX), which, according to the general mechanism proposed by Fieser and Fieser(18), would be formed at an earlier stage of the reaction. In either case, 2-hydroxy-1,4-naphthoquinone could be formed.

Further evidence relating to the structure of *isodunniol* will be found in Part II of this series.

## II. EXPERIMENTAL

All melting points are corrected. Microanalyses by Weiler and Strauss, Oxford, and by N. L. Lottkowitz of this Department.

(a) *Alkylation of 2-Hydroxy-1,4-Naphthoquinone with Trimethylallyl Bromide*

A solution of 25 g. of 2,3,3-trimethylallyl bromide(20) in 200 ml. of dry benzene was maintained at 0–5 °C., and stirred mechanically, while 40 g. of the silver salt of 2-hydroxy-1,4-naphthoquinone(21) was gradually added over a period of 75 minutes. When all the salt had been added, the mixture was stirred at the same temperature for another hour, the precipitate being thoroughly broken up when it became caked. After standing overnight, the mixture was filtered, and the solid extracted three times with warm benzene. All the filtrates were combined and extracted with 2% sodium hydroxide solution, and the residual solid was extracted with hot 5% alkali. The red alkali extracts were combined, and the benzene solution was washed with water and dried over calcium chloride.

(i) *2-Hydroxy-3-(2,3,3-Trimethylallyl)-1,4-Naphthoquinone (XIII)*.—The total red alkaline extract was carefully neutralized with acetic acid and the yellow crystalline precipitate was recrystallized from ethanol, giving silky yellow needles, m.p. 171–172 °C. Yield 12%.

Found: C, 75.1; H, 6.3%.

Calculated for  $C_{16}H_{16}O_3$ : C, 75.0; H, 6.3%.

The red filtrate remaining after the precipitation of XIII still contained the unchanged 2-hydroxy-1,4-naphthoquinone. This was precipitated by addition of hydrochloric acid. Yield, about 45%.

(ii) *2-(2,3,3-Trimethylallyloxy)-1,4-Naphthoquinone (XIV)*.—The dried benzene solution, from which the hydroxyquinones were extracted, contained the quinone ether and other products. Evaporation of the benzene under reduced pressure left a viscous, red-brown oil which readily deposited crystals on the addition of light petroleum. After washing with this solvent, the product was recrystallized from benzene-light petroleum (b.p. 40–60 °C.) with the use of charcoal. It formed long pale yellow needles, m.p. 129–130 °C.

Found: C, 74.8; H, 6.3%.

Calculated for  $C_{16}H_{16}O_3$ : C, 75.0; H, 6.3%.

(iii) *Red By-Product from the Alkylation Reaction*.—In some experiments, when the light petroleum filtrate from the precipitation of XIV was concentrated and allowed to stand in the refrigerator, a red compound was deposited. This was recrystallized from methanol, giving small orange-red plates, m.p. 145–146 °C.

Found: C, 77.6; H, 7.7%.

Calculated for  $C_{22}H_{26}O_3$ : C, 78.1; H, 7.7%.

This product appeared to be a  $\beta$ -quinone. It was insoluble in cold sodium hydroxide solution, but dissolved in aqueous sodium bisulphite to give a colourless solution from which it could be recovered by addition of alkali. The analysis suggests that it may have arisen from an impurity in the trimethylallyl bromide. It has not been investigated further.

(b) *Heterocyclic Isomers of 2-Hydroxy-3-(2,3,3-Trimethylallyl)-1,4-Naphthoquinone (XIII)*

(i)  *$\beta\gamma$ -Trimethylaldihydropyrano-1,4-Naphthoquinone (XV)*.—A solution of XIII (1.2 g.) in glacial acetic acid (10 ml.) and concentrated hydrochloric acid (4 ml.) was heated at 100 °C. for 75 minutes. The mixture was then stirred into cold water, and the precipitated oil was allowed to solidify in the refrigerator. After washing with dilute ammonia, the product was recrystallized from methanol and formed bright yellow needles, m.p. 92–93 °C.

Found: C, 74.9; H, 6.2%.

Calculated for  $C_{16}H_{16}O_3$ : C, 75.0; H, 6.3%.

(ii) *βγγ-Trimethyldihydropyrano-1,2-Naphthoquinone* (XVI).—The hydroxyquinone (XIII, 0.1 g.) was dissolved in concentrated sulphuric acid (4 ml.) and the deep red solution was allowed to stand in a stoppered flask for 20 hours. It was then poured into ice-water (15 ml.) and the orange precipitate, after washing with dilute ammonia, was crystallized from aqueous ethanol, forming glistening orange plates, m.p. 147–148 °C. The compound is soluble in aqueous sodium bisulphite, giving a colourless solution from which it is regenerated by alkali (m.p. and mixed m.p.).

Found: C, 75.0; H, 6.4%.

Calculated for  $C_{16}H_{16}O_3$ : C, 75.0; H, 6.3%.

(iii) *β-isoPropyl-β-Methyldihydropyrano-1,2-Naphthoquinone* (XVII).—Treatment of XIII (0.5 g.) with concentrated sulphuric acid (6 ml.) for only 12 minutes and then pouring on to ice gave a mixture of products. Fractional crystallization by successive additions of water to an alcoholic solution gave three fractions. The first two consisted largely of the dihydropyrano-quinone (XVI), but the third, after several crystallizations from benzene-light petroleum (b.p. 40–50 °C.), gave orange-red needles, m.p. 126–127 °C.

Found: C, 75.0; H, 6.4%.

Calculated for  $C_{16}H_{16}O_3$ : C, 75.0; H, 6.3%.

The product is considerably deeper in colour than the isomeric six-membered ring compound XVI. It appears to be a general rule that the dihydrofurano-1,2-naphthoquinones are deeper in colour than the related dihydropyrano-1,2-naphthoquinones.

From the mother liquors in the above preparation a small quantity of *βγγ-trimethyldihydropyrano-1,4-naphthoquinone* (XV) was obtained (m.p. and mixed m.p. 92–93 °C.).

(iv) *Interconversion of the Heterocyclic Isomers*.—Long treatment of the 1,4-quinone (XV) with concentrated sulphuric acid gave only the corresponding 1,2-quinone (XVI). Brief treatment produced a mixture of XVI and XVII. When a crude mixture of the 1,2-quinones (XVI and XVII), or a pure sample of XVI, was heated with concentrated hydrochloric acid for 30 minutes at 100 °C., only the single 1,4-quinone (XV) could be obtained. Identity of the products in these experiments was checked by mixed melting point determinations, and by treatment with alkali to obtain the hydroxyquinones described below.

(c) *2-Hydroxy-3-(2,3-Dimethyl-3-Hydroxybutyl)-1,4-Naphthoquinone* (XVIII)

The dihydropyrano-quinones (XV and XVI) dissolve readily on warming with aqueous alkali to give deep red solutions from which the same hydroxyhydro derivative (XVIII) may be obtained.

The finely powdered XV or XVI (1 g.) was heated at 100 °C. with sodium hydroxide solution (60 ml., 4%) for 1 hour. The mixture was then cooled, filtered, and carefully neutralized with acetic acid. The orange-yellow emulsion gradually deposited a yellow precipitate, which on recrystallization from benzene formed small yellow prisms, m.p. 135–136 °C.

Found: C, 70.4; H, 6.6%.

Calculated for  $C_{16}H_{18}O_4$ : C, 70.1; H, 6.6%.

(i) *Cyclization with Acids*.—When XVIII was heated with concentrated hydrochloric acid, the dihydropyrano-1,4-naphthoquinone (XV) was obtained (m.p. and mixed m.p. 92–93 °C.). Treatment with cold concentrated sulphuric acid, however, gave either a mixture of XVI and XVII or pure XVI, depending on the time of contact. When the deep red alkaline solution of XVIII was strongly acidified with hydrochloric acid, the dihydropyrano-1,2-naphthoquinone (XVI) slowly separated (m.p. and mixed m.p. 147–148 °C.).

(ii) *Oxidation with Alkaline Hydrogen Peroxide*.—The dihydroxyquinone (XVIII, 0.5 g.) was dissolved in sodium hydroxide solution (35 ml., 1.5%) and the red solution was boiled in a distilling flask while hydrogen peroxide (3%) was slowly added in drops. The distillate was passed into a solution of 2,4-dinitrophenylhydrazine in 5% sulphuric acid. The precipitate was recrystallized from ethanol, giving yellow needles, m.p. 124–125 °C. Mixed m.p. with acetone 2,4-dinitrophenylhydrazone, 124–125 °C., and with the 2,4-dinitrophenylhydrazone of methyl isopropyl ketone, 105–106 °C.



(d) *2-Hydroxy-3-(2,3-Dimethyl-2-Hydroxybutyl)-1,4-Naphthoquinone (XIX)*

The dihydrofurano-1,2-naphthoquinone (XVII) was warmed with 5% sodium hydroxide solution until dissolved. The deep red solution was filtered and carefully neutralized with acetic acid. A yellow crystalline precipitate separated rapidly on cooling and scratching. The dried product was crystallized twice from light petroleum (b.p. 40–50 °C.), forming small yellow leaflets, m.p. 120–121 °C.

Found: C, 70.5; H, 6.9%.

Calculated for  $C_{16}H_{18}O_4$ : C, 70.1; H, 6.6%.

This compound was also isolated, together with the isomeric dihydroxyquinone (XVIII), by fractional crystallization of the mixed products resulting from the action of alkali on the crude XVI and XVII obtained by brief treatment of XIII with concentrated sulphuric acid.

(i) *Cyclization with Acids*.—Treatment of XIX with acids gave results similar to those already described for the isomer (XVIII).

(ii) *Oxidation with Alkaline Hydrogen Peroxide*.—The oxidation of XIX was similar to that described for the isomer (XVIII) except that 1.5% hydrogen peroxide was used. The precipitated dinitrophenylhydrazone was recrystallized several times from ethanol, and then by slow evaporation of a solution in chloroform-light petroleum (b.p. 40–60 °C.). It formed yellow-orange plates, m.p. 122–123 °C.; mixed m.p. with the 2,4-dinitrophenylhydrazone of methyl isopropyl ketone, 122–123 °C., and with acetone 2,4-dinitrophenylhydrazone, 95–97 °C. The methyl isopropyl ketone for the authentic specimen was prepared by the method of Whitmore, Evers, and Rothrock(22). The 2,4-dinitrophenylhydrazone also gave no depression of melting point when mixed with the specimen obtained by Price and Robinson through chromic acid oxidation of dunnione(3).

(e) *Synthesis of Hydroxyhydroisodunnionol (III)*

The general method of Hooker(19) for the oxidation of hydroxynaphthoquinones having saturated side-chains was applied to 2-hydroxy-3-(2,3-dimethyl-3-hydroxybutyl)-1,4-naphthoquinone (XVIII).

An ice-cold solution of XVIII (0.6 g.) in sodium hydroxide (60 ml., 15%) was treated with potassium permanganate (0.6 g.) in ice-water (60 ml.). After standing for 3 hours, the precipitated manganese dioxide was removed by filtration, and the filtrate allowed to stand overnight. Careful neutralization with acetic acid gave a pale orange solution which was extracted twice with ether, and the ethereal extract was dried over sodium sulphate. Evaporation yielded a viscous residue which crystallized on manipulation with water. After two crystallizations from aqueous ethanol, the product separated from benzene-light petroleum (b.p. 40–60 °C.) in small yellow prisms, m.p. 112–113 °C. (not depressed by hydroxyhydroisodunnionol, m.p. 112–113 °C., obtained by Price and Robinson(3) from natural dunnione).

The two specimens both crystallized from aqueous ethanol as a *monohydrate*, forming small, rectangular, yellow prisms, m.p. and mixed m.p. 83 °C.

Found: Loss at 105 °C., 6.9%.

Calculated for  $C_{15}H_{16}O_4 \cdot H_2O$ : 6.5%.

The synthetic product was also cyclized to  $\alpha$ -isodunnione (V) and  $\beta$ -isodunnione (IV) by appropriate treatments with strong acids(3). The identity of these compounds was also checked by mixed melting point determinations with the products originally obtained from natural dunnione.

(f) *2-Hydroxy-3-(1,1,2-Trimethylallyl)-1,4-Naphthoquinone (XX)*

A solution of 2-(2,3,3-trimethylallyloxy)-1,4-naphthoquinone (XIV, 1 g.) in absolute ethanol (50 ml.) was refluxed for 30 minutes, during which time the colour changed from pale yellow to a deep orange-yellow. Some of the ethanol was evaporated, then addition of water caused the product to crystallize. The yield of material melting at 80–81 °C. was almost quantitative. Recrystallization from aqueous ethanol gave bright yellow needles, m.p. 82–83 °C.

Found : C, 74.8 ; H, 6.5%.

Calculated for  $C_{16}H_{16}O_3$  : C, 75.0 ; H, 6.3%.

In a similar experiment benzene was used as solvent and 60% of the ether (XIV) was recovered unchanged. There is little difference in the boiling points of the two solvents, and the concentrations and reaction times were the same. Apparently the rearrangement is much slower in benzene. This observation is significant in view of the frequent use of hydrocarbon solvents in the Claisen rearrangement(11).

(g)  $\alpha\alpha\beta\beta$ -Tetramethyldihydrofurano-1,2-Naphthoquinone (XXII)

The hydroxyquinone (XX, 0.4 g.) was dissolved in concentrated sulphuric acid (15 ml.), and after standing for 15 minutes the solution was poured into ice-water (200 ml.). The orange-red product was washed with dilute ammonia and dried. Recrystallization from light petroleum (b.p. 40–60 °C.) gave orange-red laths, m.p. 106–107 °C.

Found : C, 74.9 ; H, 6.5%.

Calculated for  $C_{16}H_{16}O_3$  : C, 75.0 ; H, 6.3%.

Oxidation with hot alkaline hydrogen peroxide gave acetone, which was identified as the 2,4-dinitrophenylhydrazone (m.p. and mixed m.p. 124–125 °C.).

(h)  $\alpha\alpha\beta\beta$ -Tetramethyldihydrofurano-1,4-Naphthoquinone (XXI)

(i) A solution of XX (0.2 g.) in concentrated hydrochloric acid (14 ml.) was boiled under reflux for 2 hours. The product was then distilled in steam, and purified further by recrystallization from light petroleum (b.p. 40–60 °C.). It formed bright yellow prisms, m.p. 107–108 °C.

Found : C, 75.2 ; H, 6.3%.

Calculated for  $C_{16}H_{16}O_3$  : C, 75.0 ; H, 6.3%.

(ii) The 1,2-quinone (XXII) was boiled under reflux with concentrated hydrochloric acid for 3 hours. The product was isolated by steam distillation and recrystallization from light petroleum as in (i), m.p. and mixed m.p. 107–108 °C.

(iii) The 1,2-quinone (XXII) was heated with 5% sodium hydroxide solution. The crystals dissolved to give a red solution which rapidly clouded and became orange in colour. On cooling, a yellow solid separated, and after washing and drying, this product crystallized from light petroleum as yellow prisms, m.p. and mixed m.p. 107–108 °C.

After standing for 15 minutes at room temperature, XXI was recovered unchanged from a solution in concentrated sulphuric acid (m.p. and mixed m.p.). When heated with sulphuric acid for 15 minutes at 100 °C., it was converted to the corresponding 1,2-quinone (XXII), m.p. and mixed m.p. 106–107 °C.

(i) 2-Hydroxy-3-(2-Methylallyl)-1,4-Naphthoquinone (XXV)

A mixture of 2-methylallyl chloride (15 g.), dry acetone (50 ml.), and sodium iodide (25 g.) was refluxed for 1 hour. The potassium salt of 2-hydroxy-1,4-naphthoquinone (20 g.) was then added, together with acetone (400 ml.) and water (100 ml.), and refluxing was continued to a total of 36 hours. The acetone was then evaporated, the residue was extracted with ether, and the ethereal solution was shaken with sodium hydroxide solution until no more hydroxyquinone was extracted. Neutralization of the red alkaline extract with acetic acid precipitated the product, which was then crystallized from ethanol (charcoal). It formed stout yellow needles, m.p. 128 °C. Yield, 17%.

Found : C, 73.0 ; H, 5.4%.

Calculated for  $C_{14}H_{12}O_3$  : C, 73.7 ; H, 5.3%.

Unchanged hydroxynaphthoquinone was precipitated from the filtrate by addition of hydrochloric acid. The ethereal solution remaining after the extraction of the hydroxyquinones contained the quinone ether. The solution was washed several times with aqueous sodium bisulphite solution and dried over sodium sulphate. Evaporation of the ether gave a yellow-brown solid which, after crystallization from benzene-light petroleum (charcoal), formed pale yellow needles.

m.p. 116–117 °C. This product was not analysed, but it was evidently 2-(2-methylallyloxy)-1,4-naphthoquinone because when heated for 30 minutes at 122–123 °C. it was rearranged to 2-hydroxy-3-(2-methylallyl)-1,4-naphthoquinone (XXV), identical with the product obtained by direct alkylation (m.p. and mixed m.p.). The latter compound was further characterized by treatment with concentrated sulphuric acid, which converts it to  $\beta\beta$ -dimethyldihydrofurano-1,2-naphthoquinone, shown by mixed m.p. to be identical with material prepared by the method of Hooker(6).

(j) *Hooker Oxidation of 2-Hydroxy-3-(2,3,3-Trimethylallyl)-1,4-Naphthoquinone (XIII)*

The general conditions of the two-stage modification by Fieser and Fieser(18) were used.

The hydroxyquinone (XIII, 0.6 g.) was added to dioxane (6 ml.), water (6 ml.), and sodium carbonate (0.3 g.). Nitrogen was passed over the solution, and the temperature was raised to 70 °C. Hydrogen peroxide (30%, 0.6 ml.) was added and heating continued for 1 hour. Then more peroxide (0.1 ml.) was added and heating continued for 30 minutes. The reaction mixture was cooled in ice, and water (2 ml.) was added. Colourless crystals separated, but the addition of hydrochloric acid gave a pale yellow emulsion from which small colourless needles separated on further cooling. This *ketol* (see Fieser and Fieser 18) was crystallized twice from benzene-light petroleum, and then from benzene, forming fine colourless needles, m.p. 177 °C. (decomp.), after softening from 155 °C.

Found: C, 66.0; H, 6.0%.

Calculated for  $C_{16}H_{18}O_5$ : C, 66.2; H, 6.2%.

This *ketol* (0.1 g.) was oxidized in the second stage by dissolving in dioxane (1.1 ml.), water (1.1 ml.), and sodium carbonate (0.07 g.), and then adding 25% sodium hydroxide (1 ml.) and copper sulphate crystals (0.5 g.) in water (2.5 ml.). The solution rapidly turned red, and after standing for 10 minutes it was warmed on the water-bath, with shaking, for a further period of 10 minutes. The cooled mixture was then filtered through Super-Cel, and the filtrate was acidified with dilute hydrochloric acid and extracted with ether. After drying and evaporating the ethereal solution, the residual yellow oil gradually solidified on chilling and rubbing. It was crystallized once from aqueous methanol by seeding, and formed golden-yellow needles, m.p. 92–95 °C. Further attempts to crystallize this material resulted in decomposition. Other preparations yielded no crystalline material at all.

The crude oily product was readily cyclized to  $\beta$ -isodunnione (IV) by the action of cold concentrated sulphuric acid, but no other resemblance to *isodunnionol* was apparent.

When XIII was oxidized with alkaline permanganate, a considerable quantity of the starting material was recovered, but no other hydroxynaphthoquinone was isolated.

(k) *Hooker Oxidation of 2-Hydroxy-3-(2-Methylallyl)-1,4-Naphthoquinone (XXV)*

Treatment of XXV with alkaline permanganate under the general conditions specified by Hooker(23) gave unchanged starting material and 2-hydroxy-1,4-naphthoquinone (m.p. and mixed m.p. 190–192 °C.). Some difficulty was also experienced with the two-stage modification of Fieser and Fieser(18) when their general directions were followed. Satisfactory degradation to the required product was finally achieved by the following procedure.

A solution of XXV (0.75 g.) in dioxane (7 ml.) and water (7 ml.) was treated with sodium carbonate (0.4 g.), and the mixture was heated to 70 °C. While nitrogen was passed over the solution, 30% hydrogen peroxide (0.8 ml.) was added, and the temperature was maintained for 40 minutes. The pale red solution was then cooled in ice, diluted with water (10 ml.) and made slightly acid with dilute hydrochloric acid. The *ketol* separated slowly as colourless needles. After 30 minutes, the product was collected and dried. When recrystallized from benzene, it formed colourless needles, m.p. 157 °C. (softening from 145 °C.).

Found: C, 63.5; H, 5.1%.

Calculated for  $C_{14}H_{14}O_5$ : C, 64.1; H, 5.4%.

This product was degraded to 2-hydroxy-3-isopropenyl-1,4-naphthoquinone (XXVI) by oxidation with alkali and copper sulphate. The *ketol* (0.35 g.) was dissolved in dioxane (4.5 ml.), and a

solution of sodium carbonate (0.25 g.) in water (4.5 ml.) was added. The solution was then treated with 25% sodium hydroxide (4 ml.) and a solution of copper sulphate crystals (1.7 g.) in water (10 ml.). The oxidation started at once and was allowed to proceed at room temperature for 10 minutes. The mixture was then heated on the water-bath, with shaking, for a further period of 10 minutes, cooled, and filtered through Super-Cel. The deep red filtrate was just acidified with dilute hydrochloric acid, and the yellow emulsion deposited light orange needles on cooling in ice. After 10 to 15 minutes, the product was collected, dried in the air, and recrystallized from light petroleum (b.p. 40–60 °C.), giving orange-yellow needles, m.p. 86 °C. Overall yield in two steps 31%.

Found: C, 73.1; H, 4.8%.

Calculated for  $C_{13}H_{10}O_3$ : C, 72.9; H, 4.7%.

(l) *Reactions of 2-Hydroxy-3-isoPropenyl-1,4-Naphthoquinone (XXVI)*

Hydrogenation of XXVI in ethanol over Adams's catalyst gave 2-hydroxy-3-isopropyl-1,4-naphthoquinone, m.p. 93–94 °C., alone or mixed with an authentic specimen prepared as described by Fieser, Hartwell, and Seligman(24).

A solution of XXVI (30 mg.) in glacial acetic acid (1 ml.) was treated with concentrated hydrochloric acid (4 ml.), and the mixture was allowed to stand in a stoppered tube for 12 hours. The crystals which had separated were recrystallized from aqueous alcohol, giving yellow prisms of 2-hydroxy-1,4-naphthoquinone, m.p. and mixed m.p. 191–192 °C. The acidic mother-liquor from which this compound separated was diluted with water and distilled into a solution of 2,4-dinitrophenylhydrazine in 5% sulphuric acid. The resulting precipitate, after recrystallization from alcohol, had m.p. 124–125 °C., alone or mixed with authentic acetone 2,4-dinitrophenylhydrazone.

### III. ACKNOWLEDGMENTS

The authors are indebted to Dr. J. R. Price for suggesting the further investigation of this problem, and for gifts of dunnione and its derivatives, to Mr. H. R. Arthur for assistance with certain preparations, to Mr. I. Parsons for assistance with the measurements of absorption spectra, and to Professor V. M. Trikojus for the use of the Beckman spectrophotometer in the Department of Biochemistry.

### IV. REFERENCES

- (1) PRICE, J. R., and ROBINSON, R.—*Nature* **142**: 147 (1938).
- (2) PRICE, J. R., and ROBINSON, R.—*J. Chem. Soc.* **1939**: 1522 (1939).
- (3) PRICE, J. R., and ROBINSON, R.—*J. Chem. Soc.* **1940**: 1493 (1940).
- (4) HOOKER, S. C.—*J. Chem. Soc.* **61**: 611 (1892).
- (5) HOOKER, S. C.—*J. Chem. Soc.* **69**: 1355 (1896).
- (6) HOOKER, S. C.—*J. Amer. Chem. Soc.* **58**: 1168 (1936).
- (7) FIESER, L. F.—*J. Amer. Chem. Soc.* **48**: 3201 (1926).
- (8) FIESER, L. F.—*J. Amer. Chem. Soc.* **49**: 857 (1927).
- (9) SMITH, L. I.—*Chem. Rev.* **27**: 287 (1940).
- (10) HAMMETT, L. P.—“Physical Organic Chemistry.” Ch. X, p. 319. (McGraw-Hill: New York and London, 1940.)
- (11) TARBELL, D. S.—“Organic Reactions.” Vol. II, Ch. I. (John Wiley & Sons, Inc.: New York, 1944.)
- (12) COOKE, R. G.—*Nature* **162**: 178 (1948).
- (13) COOKE, R. G., and SOMERS, T. C.—*Nature* **165**: 314 (1950).
- (14) HOOKER, S. C.—*J. Amer. Chem. Soc.* **58**: 1163 (1936).



- (15) COOKE, R. G., MACBETH, A. K., and WINZOR, F. L.—*J. Chem. Soc.* **1939** : 878 (1939).
- (16) MURRAY, M. J., and GALLAWAY, W. S.—*J. Amer. Chem. Soc.* **70** : 3867 (1948).
- (17) HIRSCHBERG, Y.—*J. Amer. Chem. Soc.* **71** : 3241 (1949).
- (18) FIESER, L. F., and FIESER, M.—*J. Amer. Chem. Soc.* **70** : 3215 (1948).
- (19) HOOKER, S. C.—*J. Amer. Chem. Soc.* **58** : 1174 (1936).
- (20) CLAISEN, L.—*J. prakt. Chem.* (2) **105** : 65 (1922).
- (21) FIESER, L. F.—*J. Amer. Chem. Soc.* **48** : 2922 (1926).
- (22) WHITMORE, F. C., EVERS, W. L., and ROTHROCK, H. S.—“Organic Syntheses.” Coll. Vol. II, p. 408. (John Wiley & Sons, Inc. : New York, 1943.)
- (23) HOOKER, S. C.—*J. Amer. Chem. Soc.* **58** : 1179 (1936).
- (24) FIESER, L. F., HARTWELL, J. L., and SELIGMAN, A. M.—*J. Amer. Chem. Soc.* **58** : 1223 (1936).
- (25) CRAM, D. J.—*J. Amer. Chem. Soc.* **71** : 3953 (1949).

# DUNNIONE AND RELATED NAPHTHOQUINONES

## II. SYNTHESIS OF *iso*DUNNIOL AND *dl*-DUNNIONE

By R. G. COOKE\*

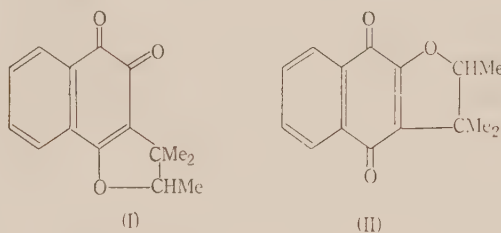
[Manuscript received March 2, 1950]

### Summary

Claisen rearrangement of 2-(3,3-dimethylallyloxy)-1,4-naphthoquinone gives two products. One of these is identical with *isodunniol*, and the other, on treatment with sulphuric acid, gives a compound which is evidently the racemic modification of the natural dextrorotatory pigment dunnione. It is shown that *isodunniol* must be 2-hydroxy-3-(1,2-dimethylallyl)-1,4-naphthoquinone, and dunnione is  $\alpha\beta$ -trimethyl-dihydrofurano-1,2-naphthoquinone. The racemic modifications of  $\alpha$ -dunnione and *allodunniol* are also described.

### I. INTRODUCTION

In Part I(1) it has been shown that the *isodunniones* are  $\alpha\beta\beta$ -trimethyl-dihydrofuranonaphthoquinones, and the establishment of their structures lends some support to the suggestion of Price and Robinson(2) that dunnione and  $\alpha$ -dunnione are represented by I and II respectively.



In amplification of a preliminary report(3), details are now given of a proof of these structures, and of the structure VI suggested(1) for *isodunniol*.

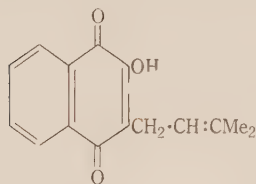
#### (a) *Synthesis of isoDunniol*

A modification of Fieser's synthesis(4) of lapachol (III) gives improved yields of this compound and of the by-product, 2-(3,3-dimethylallyloxy)-1,4-naphthoquinone (IV).

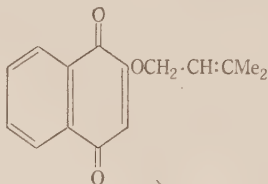
Claisen rearrangement of IV yields two compounds, which must have structures V and VI, since  $\gamma$ -substituted allyl ethers are known to rearrange with attachment at either the  $\beta$  or the  $\gamma$  carbon atom of the allyl group(5). However, there seems to be no previous record of the normal  $\gamma$ -rearrangement of a  $\gamma\gamma$ -disubstituted allyl ether. According to Claisen(6, 7),  $\gamma\gamma$ -dimethylallyl phenyl ether rearranges when heated with solid sodium carbonate, but no experimental details or proof of structure of the product are given. Rearrangement of  $\gamma\gamma$ -dimethylallyl 4-carbethoxyphenyl ether, studied by Lauer and

\* Department of Chemistry, University of Melbourne.

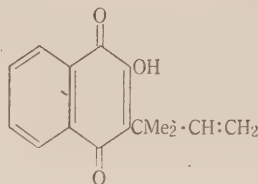
Moe(8), gives a small yield of 2,2,3-trimethyl-5-carbethoxycoumaran, which evidently results from rearrangement to the  $\beta$  carbon atom of the allyl group, followed by cyclization.



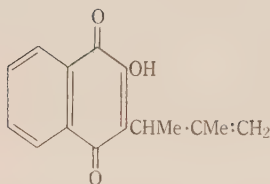
(III)



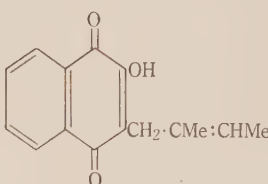
(IV)



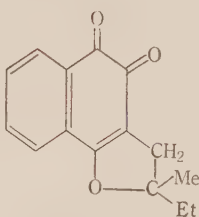
(V)



(VI)



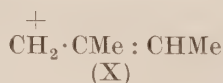
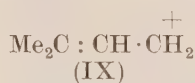
(VII)



(VIII)

Rearrangements of various substituted allyl ethers recorded in the literature(5) suggest that drastic conditions favour the abnormal reaction, in which the linkage is formed at the  $\beta$  carbon atom of the allyl group. Under the mild conditions (boiling ethanol) previously found so effective in the rearrangement of a trimethylallyloxy-1,4-naphthoquinone(1), the ether (IV) gives a practically quantitative yield of a single compound which is evidently the normal rearrangement product (V). More drastic conditions, such as boiling in xylene, or heating the ether without solvent at a temperature slightly above its melting point, give mixtures of V and VI.

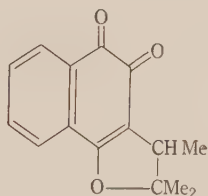
The work of Gates and Moesta(9) shows that acid catalyzed condensation of leucoisonaphthazarin with isoprene gives only two products. In accordance with theory, these are found to be lapachol (III) and the isomer (VII), the products to be expected from reaction with the carbonium ions, IX and X, formed by addition of a proton to either end of the conjugated system of isoprene.



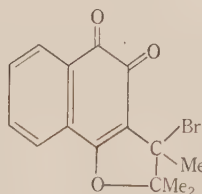
Cyclization of VII gives VIII, one of the possible structures for dunnione originally considered by Price and Robinson. Since VIII must be racemic, its melting point is not sufficient justification for the conclusion by Gates and Moesta(9) that it is not identical with dunnione. However, the synthesis of V and VI in the present investigation supports the structures assigned to VII and VIII. The preparation of III, V, VI, and VII seems to be the first instance of the occurrence of compounds having all the four possible dimethylallyl groups which can be derived from isoprene.

One of the products of Claisen rearrangement of IV is identical with *isodunniol*. This is shown by mixed melting point with the original specimen obtained from natural dunnione, by cyclization to  $\beta$ -*isodunnione* (XI), and by preparation of the *acetate* and the *leucotriacetate*. These observations, in conjunction with the results recorded in Part I(1), establish the structure VI for *isodunniol*.

It is evident that the reactions by which *isodunniol* was originally prepared are not as straightforward as Price and Robinson assumed(10). The position at which bromine attacks  $\beta$ -*isodunnione* (XI) has not been definitely determined, but as dunnione is not brominated under even more drastic conditions there seems to be no doubt that bromo- $\beta$ -*isodunnione* is correctly represented by XII.

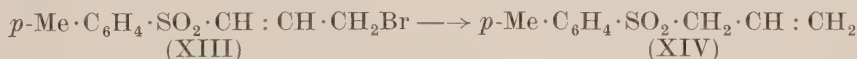


(XI)



(XII)

If this is so, debromination with zinc and alkali to give *isodunniol* (VI) must involve a rearrangement, which is in effect the migration of a double bond from the  $\alpha\beta$  to the  $\beta\gamma$  position. Analogous rearrangements occur in the course of replacement reactions in certain other systems of somewhat similar structure. According to Culvenor, Davies, and Savige(11), XIII is converted to XIV by the action of zinc and acetic acid.



(XIII)

(XIV)

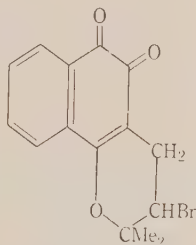
Another example is afforded by the observation of Owen(12) that XV gives XVI by reaction with potassium hydroxide in methanol



(XV)

(XVI)

The conversion of bromo- $\beta$ -lapachone (XVII) to lapachol (III) by Hooker(13) is not strictly comparable with the debromination of XI to VI, because in XVII the bromine atom is further removed from the quinone ring.



(XVII)



(b) *Synthesis of dl-Dunnione*

Treatment of 2-hydroxy-3-(1,1-dimethylallyl)-1,4-naphthoquinone (V) with concentrated sulphuric acid gives a product which is evidently the racemic modification of dunnione (I). By a coincidence, the melting point of the synthetic *dl-dunnione* (93–94 °C.) is actually quite close to that of the natural dextrorotatory pigment (98–99 °C.). Mixtures of the two melt at intermediate temperatures. On treatment with cold alkali, *dl-dunnione* undergoes a reaction similar to that of the natural dunnione, but the two products, evidently *dl- $\alpha$ -dunnione* and *dl-allodunnione*, have melting points which are lower, by 11 and 19 °C. respectively, than those of the corresponding optically active compounds described by Price and Robinson(2). The structural identity of the two pairs of compounds is shown by conversion of the *dl- $\alpha$ -dunnione* to  $\beta$ -*isodunnione* (XI) on treatment with hot concentrated sulphuric acid, and by the agreement of the absorption spectrum of *dl-allodunnione* in ethanol with that recorded by Cooke, Macbeth, and Winzor(14) for the dextrorotatory product derived from natural dunnione (see Table 1).

TABLE 1  
ABSORPTION SPECTRA (IN ETHANOL)

Compound	$\lambda_{max}$ (m $\mu$ )	log $\epsilon_{max}$	$\lambda_{max}$ (m $\mu$ )	log $\epsilon_{max}$	$\lambda_{max}$ (m $\mu$ )	log $\epsilon_{max}$
<i>dl-alloDunnione</i> ..	246	4.41	335	3.32	411	2.69
<i>d-alloDunnione</i> (ex natural dunnione)	244	4.49	334	3.34	408	2.67

Treatment of V with concentrated sulphuric acid might be expected to cause retropinacolinic rearrangement, with the ultimate formation of  $\beta$ -*isodunnione* (XI), but the cyclization to dunnione evidently occurs much more rapidly, and the work of Price and Robinson(2, 10) shows that dunnione itself is very resistant to the action of sulphuric acid, even when heated.

## II. EXPERIMENTAL

All melting points are corrected. Microanalyses by Weiler and Strauss, Oxford.

(a) *Improved Synthesis of Lapachol (III) and 2-(3,3-Dimethylallyloxy)-1,4-Naphthoquinone (IV)*

Fieser(4) reported the preparation of these two compounds in small yield by alkylation of the silver salt of 2-hydroxy-1,4-naphthoquinone with dimethylallyl bromide in absolute ether. The yields obtained were: III, 5%; IV, 3%; and recovered 2-hydroxy-1,4-naphthoquinone, 77%. The following modification of the reaction gave much better yields of both alkylation products.

The potassium salt of 2-hydroxy-1,4-naphthoquinone (43 g.) was suspended in dry acetone (320 ml.), and 3,3-dimethylallyl bromide (36 g.) was added. The mixture was boiled under reflux for 5 hours, allowed to stand overnight, and then refluxed for 7 hours. Nearly all the acetone was then distilled, and the residue was triturated with successive portions of 1% sodium hydroxide solution, each of the red extracts being removed by filtration. When no more hydroxy-naphthoquinone was extracted, the alkali insoluble residue of the quinone ether was washed with

water and dried. When recrystallized from benzene-light petroleum (b.p. 40–50 °C.), it formed pale yellow needles, m.p. 148–149 °C. Yield 12%. Fieser recorded m.p. 149–150 °C.(4).

The strongly alkaline extract containing the hydroxyquinones was carefully neutralized with acetic acid to precipitate the lapachol. Yield of material melting at 137–138 °C. was 14%. Pure lapachol melts at 140–141 °C. The filtrate from the lapachol was treated with excess hydrochloric acid to precipitate the unchanged 2-hydroxy-1,4-naphthoquinone. Yield 70%.

(b) *Claisen Rearrangement of 2-(3,3-Dimethylallyloxy)-1,4-Naphthoquinone (IV)*

(i) *Preparation of 2-Hydroxy-3-(1,1-Dimethylallyl)-1,4-Naphthoquinone (V).*—The quinone ether (IV, 4 g.) was refluxed in absolute ethanol (200 ml.) for 7 hours. The bulk of the alcohol was then evaporated at reduced pressure and the residue was treated with sodium hydroxide solution (200 ml., 0.5%). After shaking well, the mixture was filtered and the residue of unchanged ether was washed with water and dried. The dark red filtrate was acidified with acetic acid, and after standing overnight the precipitate was collected and dried. Yield of material melting at 69–70 °C. was 95% on the unrecovered quinone ether. When recrystallized from aqueous methanol, the product was obtained as yellow needles, m.p. 70–71 °C.

Found: C, 74.5; H, 5.7%.

Calculated for  $C_{15}H_{14}O_3$ : C, 74.4; H, 5.8%.

(ii) *Preparation of 2-Hydroxy-3-(1,2-Dimethylallyl)-1,4-Naphthoquinone (isodunnionol, VI).*—A solution of IV (1 g.) in xylene (50 ml.) was refluxed for 45 minutes. The cooled solution was then extracted with sodium hydroxide solution, and the deep red alkaline extract was just neutralized with acetic acid. When precipitation was complete, the mixed products were separated by crystallization from ethanol, the abnormal rearrangement product (VI) being readily isolated as the less soluble isomer. It was obtained as yellow needles, m.p. 118–119 °C., alone or mixed with the authentic specimen of *isodunnionol* originally obtained from natural dunnione(10). Yield 20%.

The *acetate* was prepared by treatment with acetic anhydride and pyridine at room temperature for 30 minutes. Crystallized from aqueous ethanol, it formed pale yellow needles, m.p. 73–74 °C. The same compound (m.p. and mixed m.p.) was obtained from authentic *isodunnionol*.

Found: C, 71.7; H, 5.7%.

Calculated for  $C_{17}H_{16}O_4$ : C, 71.8; H, 5.7%.

The *leucotriacetate* was prepared by shaking VI with acetic anhydride, pyridine, and zinc dust until the solution became colourless. The mixture was then filtered into dilute hydrochloric acid and stirred until the product crystallized. Recrystallization from aqueous acetone gave colourless prisms, m.p. 124–126 °C. The same product was obtained (m.p. and mixed m.p.) by reductive acetylation of authentic *isodunnionol*.

Found: C, 67.8; H, 6.3%.

Calculated for  $C_{21}H_{22}O_6$ : C, 68.1; H, 6.0%.

Treatment of the synthetic *isodunnionol* with concentrated sulphuric acid for 24 hours at room temperature gave  $\beta$ -*isodunnione*, m.p. 129–131 °C., alone or mixed with a specimen derived from natural dunnione.

After crystallization of the synthetic *isodunnionol*, the alcoholic mother-liquors yielded the normal rearrangement product (V) on dilution with water. Further crystallization from aqueous methanol gave yellow needles, m.p. 70–71 °C., alone or mixed with the product obtained as in (i). Yield about 70%.

(iii) The same two rearrangement products were obtained, in lower yield, when the ether (IV, 0.3 g.) was heated, without solvent, for 20 minutes at 153 °C. Only *isodunnionol* was obtained pure, the normal rearrangement product being obtained in only small quantity, m.p. 66–68 °C.

(c) *dl-Dunnione*

2-Hydroxy-3-(1,1-dimethylallyl)-1,4-naphthoquinone (V, 3 g.) was dissolved in concentrated sulphuric acid (30 ml.) and the solution allowed to stand for 15 minutes. It was then poured

slowly into ice and water with vigorous stirring, and after standing overnight the product was recrystallized from light petroleum (40–50 °C.). It was obtained as bright orange-red, striated laths, m.p. 93–94 °C. Yield 86%. A mixture with natural *d*-dunnione (m.p. 98–99 °C.) melted at 93–97 °C.

Found: C, 74.6; H, 6.0%.

Calculated for  $C_{15}H_{14}O_3$ : C, 74.4; H, 5.8%.

(d) *Action of Alkali on Synthetic dl-Dunnione*

(i) *dl-α-Dunnione*.—Synthetic *dl*-dunnione (0.2 g.) was immersed in sodium hydroxide solution (50 ml., 5%) and allowed to stand for 17 hours. The orange-red crystals had then been replaced by a yellow solid, and the alkaline solution was also yellow. The dried solid was recrystallized from light petroleum (40–50 °C.), giving yellow plates, m.p. 110–111 °C. Price and Robinson reported m.p. 121–122 °C. for *d-α*-dunnione(2).

Found: C, 74.7; H, 5.8%.

Calculated for  $C_{15}H_{14}O_3$ : C, 74.4; H, 5.8%.

Treatment with concentrated sulphuric acid at 100 °C. for 30 minutes converted this compound into *β*-isodunnione identical with that obtained by similar treatment of the dextro-rotatory *α*-dunnione obtained from natural dunnione. (M.p. and mixed m.p. 129–131 °C.)

(ii) *dl-alloDunnione*.—The yellow alkaline filtrate from the *dl-α*-dunnione was acidified with hydrochloric acid, and after standing for a few hours the precipitate was recrystallized from ethanol as yellow plates, m.p. 142–143 °C. Price and Robinson reported m.p. 161–162 °C. for *d-allo*dunnione(2).

Found: C, 74.4; H, 6.0%.

Calculated for  $C_{15}H_{14}O_3$ : C, 74.4; H, 5.8%.

This compound is also obtained as the sole product of the action of *hot* alkali on synthetic *dl*-dunnione and *dl-α*-dunnione, under the conditions previously applied to the corresponding *dextro* isomers(2).

### III. ACKNOWLEDGMENT

The author is indebted to Mr. R. D. Brown for the measurement of the ultraviolet absorption spectrum of *dl-allo*dunnione.

### IV. REFERENCES

- (1) COOKE, R. G., and SOMERS, T. C.—*Aust. J. Sci. Res. A* **3**: 466 (1950).
- (2) PRICE, J. R., and ROBINSON, R.—*J. Chem. Soc.* **1939**: 1522 (1939).
- (3) COOKE, R. G.—*Nature* **162**: 178 (1948).
- (4) FIESER, L. F.—*J. Amer. Chem. Soc.* **49**: 857 (1927).
- (5) TARBELL, D. S.—“Organic Reactions.” Vol. II, Ch. I. (John Wiley & Sons, Inc.: New York, 1944.)
- (6) CLAISEN, L.—*J. prakt. Chem.* (2) **105**: 65 (1922).
- (7) CLAISEN, L., and TIETZE, E.—*Ber. dtsh. chem. Ges.* **59**: 2344 (1926).
- (8) LAUER, W. M., and MOE, O.—*J. Amer. Chem. Soc.* **65**: 289 (1943).
- (9) GATES, M., and MOESTA, D. L.—*J. Amer. Chem. Soc.* **70**: 614 (1948).
- (10) PRICE, J. R., and ROBINSON, R.—*J. Chem. Soc.* **1940**: 1493 (1940).
- (11) CULVENOR, C. C. J., DAVIES, W., and SAVIGE, W. E.—*J. Chem. Soc.* **1949**: 2198 (1949).
- (12) OWEN, L. N.—*J. Chem. Soc.* **1949**: 236 (1949).
- (13) HOOKER, S. C.—*J. Chem. Soc.* **61**: 611 (1892).
- (14) COOKE, R. G., MACBETH, A. K., and WINZOR, F. L.—*J. Chem. Soc.* **1939**: 878 (1939).

# DUNNIONE AND RELATED NAPHTHOQUINONES

## III. REARRANGEMENT OF HYDROXYNAPHTHOQUINONES TO INDENONE CARBOXYLIC ACIDS. *alloodunnione*

By R. G. COOKE\* and T. C. SOMERS\*

[*Manuscript received March 2, 1950*]

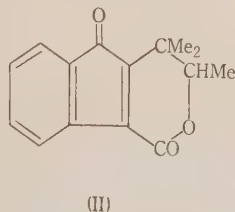
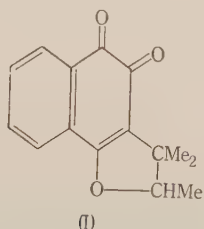
### Summary

Certain *tert.*-alkylhydroxynaphthoquinones are found to isomerize rapidly in hot aqueous alkali, and the reaction is correlated with the rearrangement of dunnione to *alloodunnione*. The change also occurs, relatively slowly, with 2-hydroxy-3-*isopropyl*-1,4-naphthoquinone. An alternative synthesis shows that the product in this case is 2-*isopropyl*-indenone-3-carboxylic acid. The ethyl ester of this acid shows an absorption spectrum very similar to that of *alloodunnione*.

Alkaline rearrangement of 2-hydroxy-3-(1,1-dimethylallyl)-1,4-naphthoquinone evidently gives the corresponding dimethylallylindenonecarboxylic acid. Treatment of the latter with concentrated sulphuric acid forms *alloodunnione*, and a new isomer which apparently results from a retropinacolinic rearrangement. All these observations support the structure originally proposed for *alloodunnione*.

### I. INTRODUCTION

It was shown by Price and Robinson(1, 2) that the natural pigment, dunnione (I), undergoes a remarkable rearrangement when heated with aqueous alkali. A red solution is first formed, but the colour changes to yellow in the course of a few minutes, and mineral acid then precipitates a yellow isomer called *alloodunnione*. Although this compound resembles the heterocyclic naphthoquinones in some of its chemical reactions, its absorption spectrum is unlike those of the naphthoquinones(3). The structure (II) tentatively suggested by Price and Robinson(1) seems consistent with most of their observations, but until recently there has been no strong evidence to support it. Although ring contraction of naphthoquinones to indane derivatives is not uncommon(4, 5, 6), no compounds closely related to II were known.



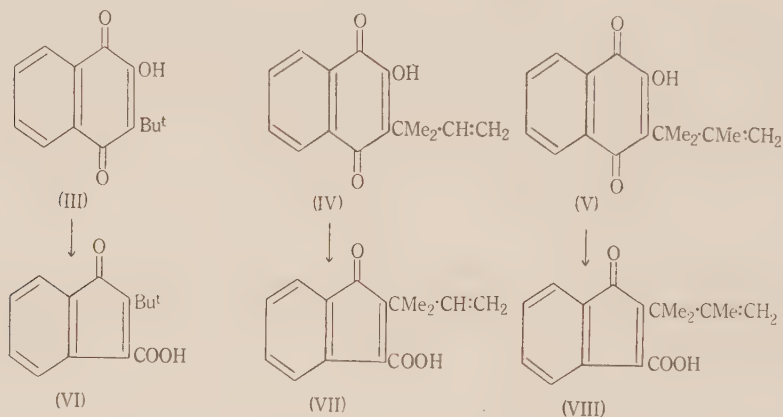
\* Department of Chemistry, University of Melbourne.



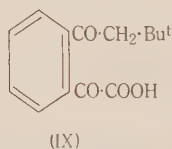
However, other examples of this type of rearrangement have been observed recently(7-10), and the contributions from this laboratory are now presented in detail.

(a) *Rearrangement of Hydroxynaphthoquinones*

The establishment of the structure of dunnione, and the observation that synthetic *dl*-dunnione readily rearranges to *dl*-allodunnione(11) suggested that the rapid isomerization in alkali might be a common property of hydroxynaphthoquinones having tertiary alkyl substituents. Investigation of three such compounds, viz. 2-hydroxy-3-*tert.*-butyl-1,4-naphthoquinone (III), 2-hydroxy-3-(1,1-dimethylallyl)-1,4-naphthoquinone (IV), and 2-hydroxy-3-(1,1,2-trimethylallyl)-1,4-naphthoquinone (V), shows that rearrangement to isomeric *acids* occurs rapidly in boiling aqueous alkali. The products are all yellow crystalline compounds which dissolve in aqueous sodium bicarbonate, giving *yellow* solutions. They are therefore not hydroxynaphthoquinones, and on the basis of the structure (II) suggested for *allodunnione* they may be represented by formulae VI, VII, and VIII respectively.



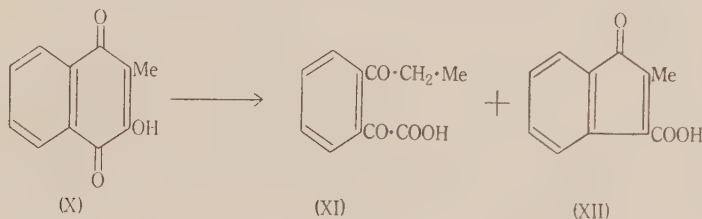
Subsequently it was learned that the rearrangement of the *tert.*-butyl-hydroxynaphthoquinone (III) had been observed independently by Dr. M. G. Ettlinger of Harvard (personal communication). Ettlinger also found that treatment of III with hot aqueous disodium hydrogen phosphate gives a colourless intermediate compound, believed to be IX, which is converted to VI by hot alkali.



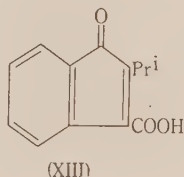
Earlier work by Fieser(12) shows that glyoxylic acid derivatives analogous to IX are obtainable by alkaline cleavage of certain hydroxyphenanthrenequinones.

Further investigations show that the rearrangement of hydroxynaphthoquinones is not confined to compounds having a tertiary alkyl substituent.

According to Fieser(8), 2-hydroxy-3-*cyclohexyl*-1,4-naphthoquinone is also isomerized to a yellow acid by boiling aqueous alkali, while Shemyakin and co-workers(9) find that phthiocol (X) is converted to two products (apparently XI and XII) on long heating in weakly alkaline buffers.



Finally, in this laboratory, it is found that 2-hydroxy-3-*isopropyl*-1,4-naphthoquinone rearranges to 2-*isopropylinden*-1-one-3-*carboxylic acid* (XIII).



The structure of XIII is proven by the synthesis outlined in Figure 1. This synthesis provides the first direct evidence that all these rearrangement products are derivatives of indenone-3-carboxylic acid.

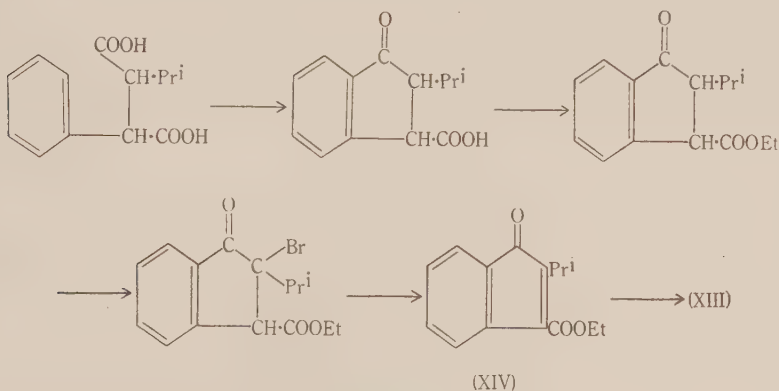
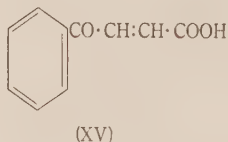


Fig. 1

As postulated by Shemyakin and co-workers(9), the rearrangement of the naphthoquinones evidently involves two successive steps: firstly, hydrolytic fission of the quinone ring with the formation of an acyl phenylglyoxylic acid, and secondly, cyclization of this intermediate by an ordinary Claisen type condensation. The first of these steps, as pointed out by Fieser(12), bears

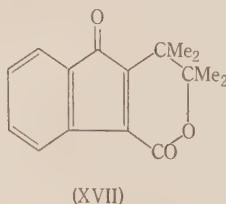
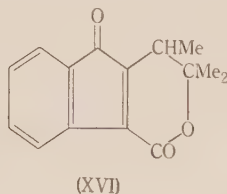
a superficial resemblance to the alkaline hydrolysis of simple  $\beta$ -diketones. However, compounds like  $\beta$ -benzoylacrylic acid (XV) undergo similar hydrolysis in alkali.



Apparently this first step is favoured by the attachment of a tertiary alkyl group to the quinone ring. Compounds having such groups rearrange in a few minutes, but with compounds having secondary or primary alkyl groups the reaction is relatively slow. The activating effect of the tertiary groups is not yet clear.

(b) *Structure of alloDunnione*

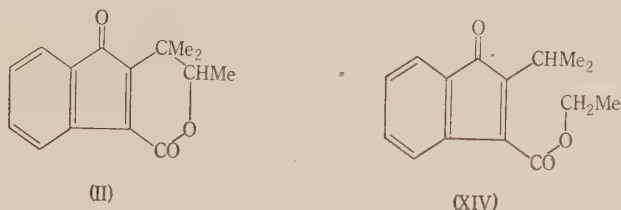
From the results of oxidation reactions Price and Robinson(1, 2) concluded that the isoprenoid chain is still intact in *allodunnione*. Thus, oxidation with alkaline hydrogen peroxide gives acetaldehyde and an acid  $C_{12}H_{12}O_4$ , while chromic acid forms acetone. Further evidence on this point is provided by a study of the action of strong acids on VII and VIII. Treatment of VII with strong hydrochloric acid gives *dl-allodunnione*. The same compound is obtained by brief treatment of VII with concentrated sulphuric acid, but prolonged standing in the latter reagent produces a new isomer. As this product gives acetone on oxidation with alkaline hydrogen peroxide, it evidently results from a retropinacolinic rearrangement of an intermediate carbonium ion, and the structure XVI is assigned to it. Evidently the lactone ring is not closed until the sulphuric acid solution is diluted. Once it is formed, *allodunnione* is not changed by long standing in concentrated sulphuric acid.



The action of strong acids on VIII produces only a single compound, which is evidently the *allodunnione* homologue (XVII). This has not been obtained by the action of alkali on the tetramethyldihydrofuranonaphthoquinones which are formed by cyclization of V(13).

The structure (II) for *allodunnione* is further supported by a comparison of the light absorption of this compound with that of 2-isopropyl-3-carbethoxy-indenone (XIV), the synthesis of which was designed to provide an open model of

structure II. The close resemblance of the two structures is emphasized by writing them as shown below.



Although the spatial arrangement of the substituents in XIV would not actually correspond to the lactone ring structure in II, the spectra of the two compounds should be very similar.

The absorption spectrum of XIV in ethanol was measured with a Beckman spectrophotometer (Model DU). The maxima are recorded in Table 1, together with the figures for *allodunnione* previously recorded by Cooke, Macbeth, and Winzor(3).

TABLE 1

Compound	$\lambda$ (m $\mu$ )	$\log \epsilon$	$\lambda$ (m $\mu$ )	$\log \epsilon$	$\lambda$ (m $\mu$ )	$\log \epsilon$
<i>iso</i> Propylcarbethoxyindenone (XIV) ..	244	4.43	328	3.13	405	2.65
<i>allo</i> Dunnione .. ..	244	4.49	334	3.34	408	2.67

These results, together with the other observations recorded in this series of papers, lead to the conclusion that *allodunnione* is correctly represented by structure II.

## II. EXPERIMENTAL

All melting points are corrected unless otherwise stated. Microanalyses by Weiler and Strauss, Oxford, and by N. L. Lottkowitz of this Department.

### (a) Preparation of 2-Hydroxy-3-tert.-Butyl-1,4-Naphthoquinone (III)

A mixture of 0.6 g. of 2-*tert.*-butyl-1,4-naphthoquinone(14) with glacial acetic acid (5 ml.) and anhydrous sodium acetate (11 g.) was treated with a solution of bromine (0.45 g.) in glacial acetic acid (2 ml.). The red solution was allowed to stand in a stoppered flask at room temperature for 8 days. Water (10 ml.) was then added, and the crude bromo compound which precipitated was collected and dried. The product (0.6 g.) was hydrolysed by dissolving in methanol (25 ml.) and adding sodium hydroxide pellets (0.6 g.). After heating on the steam-bath for 1 hour, the deep red solution was cooled, filtered, and acidified with hydrochloric acid. The product was extracted with ether, and after washing with a little dilute sodium carbonate solution the extract was dried and evaporated. The solid yellow residue (0.4 g.) was recrystallized, first from aqueous methanol and then from light petroleum (b.p. 40–60 °C.). It formed silky yellow needles, m.p. 91 °C.



The compound was also obtained in small yield by alkylation of 2-hydroxy-1,4-naphthoquinone with pivalic acid and red lead, using acetoacetic ester as promoter. This general method of Fieser and Chang(15) has not previously been used with hydroxynaphthoquinone. The product was isolated by steam distillation and crystallization from aqueous methanol. Yield <2%.

After this work was completed, Fieser and co-workers(16) described the preparation of this compound (m.p. 92.5 °C.) by a third method, viz. Hooker oxidation of the higher homologue.

(b) *Preparation of 2-Hydroxy-3-isoPropyl-1,4-Naphthoquinone*

This compound was prepared most conveniently by the method of Fieser, Hartwell, and Seligman(17), but it was also obtained by alkylation of hydroxynaphthoquinone with *isobutyric* acid and red lead, using acetoacetic ester as promoter (compare preparation of *tert.*-butyl derivative above). The product was isolated by steam distillation, and after one crystallization from aqueous ethanol it had m.p. 90–92 °C. Yield approx. 1.5%.

(c) *Rearrangement of Hydroxynaphthoquinones in Alkali*

(i) *2-Hydroxy-3-tert.-Butyl-1,4-Naphthoquinone*.—The quinone (0.1 g.) was dissolved in aqueous sodium hydroxide (4%, 15 ml.), and the solution was boiled for 10 minutes. The colour had then faded to pale orange, and after clarification with charcoal the addition of hydrochloric acid precipitated a yellow crystalline product. After crystallization from light petroleum (b.p. 40–60 °C.), the *acid* was obtained as large glistening yellow plates, m.p. 165 °C. It dissolves in aqueous sodium bicarbonate, giving an orange-yellow solution.

Found: C, 73.4; H, 5.7%.

Calculated for  $C_{14}H_{14}O_3$ : C, 73.1; H, 6.1%.

(ii) *2-Hydroxy-3-(1,1-Dimethylallyl)-1,4-Naphthoquinone (11)*.—A solution of the quinone (0.5 g.) in 5% sodium hydroxide (50 ml.) was refluxed for 1 hour. The cooled orange solution was acidified with hydrochloric acid, and the yellow crystalline *acid* was crystallized from benzene-light petroleum, giving bright yellow leaflets, m.p. 152.5 °C. Yield 90%.

Found: C, 74.4; H, 5.9%.

Calculated for  $C_{15}H_{14}O_3$ : C, 74.4; H, 5.8%.

(iii) *2-Hydroxy-3-(1,1,2-Trimethylallyl)-1,4-Naphthoquinone (13)*.—Rearrangement of V was complete after refluxing for 2 hours in 5% aqueous sodium hydroxide. After extraction with ether, the solution was acidified and the *product* recrystallized from light petroleum. It formed rosettes of small bright yellow needles, m.p. 123.5–124.5 °C. Yield 90%.

Found: C, 74.7; H, 6.1%.

Calculated for  $C_{16}H_{16}O_3$ : C, 75.0; H, 6.3%.

(iv) *2-Hydroxy-3-isopropyl-1,4-Naphthoquinone*.—The rearrangement of this quinone does not proceed as smoothly as with the tertiary alkyl compounds. The solution darkens whether the reaction is conducted under ordinary reflux conditions, in an open beaker with free access of air, or in a nitrogen atmosphere. A blue salt, soluble in water and in ether, is formed as a by-product. A similar substance was observed in the rearrangement of 2-hydroxy-3-cyclohexyl-1,4-naphthoquinone by Fieser(8). The best results were finally obtained by the following procedure. The quinone (0.2 g.) was dissolved in aqueous sodium hydroxide (5%, 30 ml.) and the solution was refluxed for 2½ hours. The brownish-red mixture was cooled and extracted several times with ether to remove the blue salt. The aqueous layer was then acidified and again extracted with ether. The ether extract was shaken three times with aqueous sodium bicarbonate, and the orange aqueous solution so obtained gave a yellow crystalline precipitate on treatment with hydrochloric acid. (Some unchanged quinone was found in the ether layer.) After recrystallization from light petroleum, the *acid* formed silky yellow needles, m.p. 146–147 °C. Yield 70%.

Found: C, 72.0; H, 5.5%.

Calculated for  $C_{13}H_{12}O_3$ : C, 72.2; H, 5.6%.

(d) *Synthesis of 2-isoPropylindenone-3-Carboxylic Acid (XIII)*

(i) *Phenylisopropylsuccinic Acid*.—Phenylisopropylsuccindinitrile was prepared by the method of Upson and Thompson(18), who recorded that the compound was difficult to hydrolyse to the dibasic acid. Various methods of hydrolysis have now been tried in an effort to improve upon the results of these workers.

The dinitrile (5 g.) was dissolved in concentrated sulphuric acid (35 ml.) and the solution was heated on the steam-bath for 1 hour, then cooled and poured into ice-water (60 ml.). The white precipitate was collected on a sintered glass funnel and recrystallized from a large volume of methanol, forming microscopic hexagonal crystals which begin to decompose at about 300 °C., and melt at 330 °C. (uncorr.). The compound, evidently *phenylisopropylsuccindiamide*, is practically insoluble in most organic solvents and insoluble in sodium hydroxide.

Found: N, 12.0%.

Calculated for  $C_{13}H_{18}O_2N_2$ : N, 12.0%.

When boiled with aqueous alcoholic sodium hydroxide, ammonia was evolved, and acidification of the solution then liberated phenylisopropylsuccinic acid, which was extracted with benzene and, after evaporation of the solvent, was finally crystallized from water, forming small colourless prisms, m.p. 178–179 °C. Upson and Thompson recorded m.p. 178 °C.(18).

The diamide was also hydrolysed by refluxing in glacial acetic acid-hydrochloric acid mixture for 25 hours. Some phenylisopropylsuccinic acid was isolated, but the hydrolysis did not appear to be complete. These methods were inconvenient for the production of the dibasic acid in quantity, but it was found that refluxing the diamide with sulphuric acid (50% w/w) caused rapid hydrolysis and partial cyclodehydration of the dibasic acid. As the further treatment of the mixed acids with concentrated sulphuric acid completed the formation of the indanone acid, this method was used for the preparation of the latter.

(ii) *2-isoPropylindanone-3-Carboxylic Acid*.—Phenylisopropylsuccindinitrile (10.4 g.) was dissolved in concentrated sulphuric acid (74 ml.) and the solution was heated on the steam-bath while water (137 ml.) was added slowly with shaking. A white precipitate separated, and the mixture was then refluxed for 10 hours, during which time the precipitated powder became crystalline. After cooling, the solid was collected on a sintered glass funnel, washed with water, and dried in a steam oven. Yield 11.6 g., colourless needles and plates, m.p. 160–178 °C. The product was free of nitrogen.

The crude mixture of acids was finely powdered and heated with concentrated sulphuric acid (35 ml.) on the steam-bath for  $2\frac{1}{2}$  hours. After thorough cooling, the mixture was run slowly into ice-water (140 ml.). Seed was first obtained from a few drops of the mixture, so that the rest of the product crystallized readily during the dilution. After thorough cooling in the refrigerator, the product was recovered, dried, and recrystallized from benzene-light petroleum. (Extraction of the aqueous filtrate with benzene recovered a further quantity of the product.) It formed rosettes of colourless needles, m.p. 103–104 °C. Yield 71% from the dinitrile.

Found: C, 71.2; H, 6.7%.

Calculated for  $C_{13}H_{14}O_3$ : C, 71.5; H, 6.4%.

(iii) *2-isoPropyl-3-Carbethoxyindanone*.—The acid was esterified with ethanol and dry hydrogen chloride, the mixture being refluxed for 4 hours. After drying in ether, the recovered ester was crystallized from light petroleum, giving large colourless prisms, m.p. 35–36 °C. Yield 95%.

Found: C, 73.5; H, 7.3%.

Calculated for  $C_{15}H_{18}O_3$ : C, 73.2; H, 7.3%.

The semicarbazone of the ester separated from aqueous alcohol in stout needles, m.p. 226 °C. (decomp.).

Found: N, 13.9%.

Calculated for  $C_{16}H_{21}O_3N_3$ : N, 13.9%.

(iv) *2-isoPropyl-3-Carbethoxyindenone (XIV)*.—The saturated ester (7.65 g.) was dissolved in carbon tetrachloride (15 ml.) and treated with bromine (5.0 g.) in carbon tetrachloride (15 ml.).

The reaction was initiated by adding a few drops of the bromine solution and heating the mixture at 45 °C. until the reaction started. The mixture was then cooled to 20 °C., and the rest of the bromine solution was added in drops, with stirring, over a period of 30 minutes. Stirring was continued for a further 30 minutes, the solution then being washed once with water (10 ml.), dried, and evaporated under reduced pressure. The residue was a nearly colourless viscous oil (10.1 g.) which could not be crystallized, and was not purified further.

Found: Br, 25.5%.

Calculated for  $C_{15}H_{17}O_3Br$ : Br, 24.6%.

The crude bromo-ester (10 g.) was dissolved in freshly distilled diethylaniline (60 ml.) and the solution was refluxed gently for 1 hour. After cooling, excess dilute sulphuric acid was added and the mixture was extracted with ether. The ether layer was washed successively with dilute acid, sodium hydroxide, and water. After drying, the ether was evaporated and the residual oil was fractionated under reduced pressure. The unsaturated ester was obtained as a viscous, orange oil, b.p. 140–143 °C./0.5 mm. Yield 82%. For analysis and the measurement of absorption spectrum the compound was further purified by passing a solution in benzene over a column of activated alumina. The benzene was evaporated from the eluate, and the residue was again fractionated under reduced pressure. The whole distilled over the range 146–148 °C./0.75 mm. A middle fraction, b.p. 146.5–147 °C./0.75 mm., was used for analysis and spectroscopy. It was a bright orange oil, readily soluble in ethanol, benzene, and light petroleum. Below 0 °C. it can be solidified to yellow-orange needles.

Found: C, 73.9; H, 6.8%.

Calculated for  $C_{15}H_{16}O_3$ : C, 73.8; H, 6.6%.

(v) *2-isopropylindenone-3-Carboxylic Acid (XIII)*.—The above ester (0.4 g.) was dissolved in ethanol (7 ml.) and the solution was heated to boiling. Sodium hydroxide solution (5%, 10 ml.) was then added, and the mixture was heated on the steam-bath, with shaking, for a further period of 3 minutes. The cooled mixture was acidified with dilute hydrochloric acid and extracted with ether, the ether layer then being extracted four times with 10 ml. portions of 5% sodium hydroxide solution. Acidification of the combined alkaline extracts gave a yellow crystalline precipitate. After recrystallization from light petroleum, the product formed silky yellow needles, m.p. 146–147 °C., alone or mixed with the acid obtained by rearrangement of 2-hydroxy-3-isopropyl-1,4-naphthoquinone. Yield 90%.

The *methyl ester* was prepared by esterification with methanol and dry hydrogen chloride. It readily solidified and was recrystallized from aqueous ethanol, giving yellow-orange needles, m.p. 46–47 °C.

Found: C, 72.9; H, 6.1; OMe, 13.5%.

Calculated for  $C_{14}H_{14}O_3$ : C, 73.1; H, 6.1; OMe, 13.5%.

#### (e) Cyclization of the Acids VII and VIII

(i) *Action of Hydrochloric Acid on VII*.—A solution of VII (20 mg.) in aqueous sodium hydroxide (5%, 0.5 ml.) was strongly acidified with concentrated hydrochloric acid (10 ml.). The yellow emulsion deposited orange-yellow needles in the course of 5 days. The product was extracted with ether, the ethereal solution was washed with 2% alkali, and after evaporation of the ether the residue was recrystallized from light petroleum (b.p. 40–60 °C.), giving orange-yellow prisms, m.p. and mixed m.p. with *dl*-allodunnione(11) 142 °C.

(ii) *Action of Sulphuric Acid on VII*.—Treatment of VII (0.1 g.) with concentrated sulphuric acid (3 ml.) gave a brown-red solution. After 1 minute this was poured on to ice, and the mixture allowed to stand for 1 hour. The product was extracted with ether, recovered, and recrystallized from light petroleum, giving orange-yellow prisms, m.p. 142–143 °C. alone or mixed with *dl*-allo-dunnione(11).

In another experiment the sulphuric acid solution of VII was allowed to stand in a stoppered flask for 24 hours before pouring on ice. The mixture was then extracted with ether, and the ethereal solution was washed with 2% alkali until this was no longer coloured blue. The product

recovered from the ether was then dissolved in light petroleum and passed over a column of activated alumina. The eluate was evaporated and the residue was recrystallized from aqueous acetone, giving yellow-orange rectangular plates, m.p. 117–118 °C.

Found: C, 73.9; H, 5.8%.

Calculated for  $C_{15}H_{14}O_3$ : C, 74.4; H, 5.8%.

This compound is evidently the *lactone* (XVI). It dissolves when heated with aqueous alkali to give a yellow solution from which it is recovered on acidification, m.p. and mixed m.p. 117–118 °C. When oxidized with boiling alkaline hydrogen peroxide, the distilled vapours contained acetone, which was identified as the 2,4-dinitrophenylhydrazone (m.p. and mixed m.p.). Acidification of the alkaline mixture gave an acid, m.p. 125–126 °C.

(iii) Treatment of VIII with concentrated sulphuric acid, or with somewhat diluted hydrochloric acid, gave only one product which was recrystallized from aqueous ethanol, forming yellow-orange needles, m.p. 111–112 °C.

Found: C, 75.0; H, 6.3%.

Calculated for  $C_{16}H_{16}O_3$ : C, 75.0; H, 6.3%.

This *lactone* (XVII) dissolves in hot aqueous alkali, giving a yellow solution, from which it is reprecipitated by acids. Oxidation with hot alkaline hydrogen peroxide gives acetone (identified as the 2,4-dinitrophenylhydrazone, m.p. and mixed m.p. 125–126 °C.) and an acid which crystallized from water in long colourless needles, m.p. 125–126 °C. not depressed by mixing with the acid similarly obtained from XVI. This acid was analysed, but has not been examined further.

Found: C, 69.6, 69.3; H, 6.7, 6.6%.

#### (f) Action of Aqueous-Alcoholic Alkali on Tetramethyldihydrofurano-1,2-Naphthoquinone

It was shown previously(13) that aqueous alkali converted this dunnione homologue to the corresponding 1,4-quinone, but did not cause a rearrangement analogous to the conversion of dunnione to *allodunnione*, evidently because the dihydrofurano ring did not open readily in alkali. When aqueous-alcoholic alkali is used, the ring apparently does open, and a red solution is formed. However, the compound XVII is not formed on boiling. A product, not isomeric with the starting material, was isolated in small yield. It crystallized from light petroleum in stout yellow needles, m.p. 210 °C. (decomp.). It forms an almost colourless solution when heated with alkali, but is regenerated on acidification.

Found: C, 72.4; H, 5.2%.

Calculated for  $C_{13}H_{11}O_3$ : C, 72.5; H, 5.1%.

### III. ACKNOWLEDGMENTS

The authors are indebted to Dr. J. R. Price for his continued interest in this work, and for helpful discussions, to Mr. I. Parsons for assistance with the measurement of the absorption spectrum, and to Dr. M. G. Ettlinger and Professor L. F. Fieser for advance information on work done at Harvard.

### IV. REFERENCES

- (1) PRICE, J. R., and ROBINSON, R.—*J. Chem. Soc.* **1939**: 1522 (1939).
- (2) PRICE, J. R., and ROBINSON, R.—*J. Chem. Soc.* **1940**: 1493 (1940).
- (3) COOKE, R. G., MACBETH, A. K., and WINZOR, F. L.—*J. Chem. Soc.* **1939**: 878 (1939).
- (4) ZINCKE, T.—*Ber. deutsch. chem. Ges.* **21**: 1027, 2379 (1888).
- (5) SCHMIDT, J.—*Ber. deutsch. chem. Ges.* **33**: 543 (1900).
- (6) UFIMTSEV, V. N.—*C.R. Acad. Sci. U.R.S.S.* **51**: 517 (1946).
- (7) COOKE, R. G.—*Nature* **162**: 178 (1948).
- (8) FIESER, L. F.—*J. Amer. Chem. Soc.* **70**: 3237 (1948).
- (9) SHCHUKINA, L. A., KONDRAT'eva, A. P., and SHEMAKIN, M. M.—*J. Gen. Chem.* **18**: 2121 (1948).
- (10) COOKE, R. G., and SOMERS, T. C.—*Nature* **165**: 314 (1950).



- (11) COOKE, R. G.—*Aust. J. Sci. Res. A* **3** : 481 (1950).
- (12) FIESER, L. F.—*J. Amer. Chem. Soc.* **51** : 940, 1896 (1929).
- (13) COOKE, R. G., and SOMERS, T. C.—*Aust. J. Sci. Res. A* **3** : 466 (1950).
- (14) BROMBY, N. G., PETERS, A. T., and ROWE, F. M.—*J. Chem. Soc.* **1943** : 144 (1943).
- (15) FIESER, L. F., and CHANG, F. C.—*J. Amer. Chem. Soc.* **64** : 2043 (1942).
- (16) FIESER, L. F., ET AL.—*J. Amer. Chem. Soc.* **70** : 3174 (1948).
- (17) FIESER, L. F., HARTWELL, J. L., and SELIGMAN, A. M.—*J. Amer. Chem. Soc.* **58** : 1223 (1936).
- (18) UPSON, F. W., and THOMPSON, T. J.—*J. Amer. Chem. Soc.* **44** : 181 (1922).

# THE SYNTHESIS OF MELICOPICINE AND SOME TRIMETHOXY-10-METHYLACRIDONES

By G. K. HUGHES,\* K. G. NEILL,\* and E. RITCHIE\*

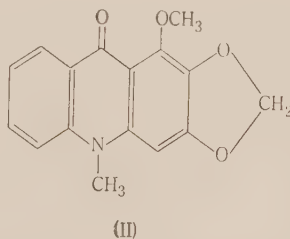
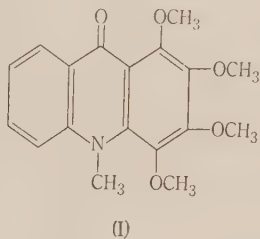
[Manuscript received May 29, 1950]

## Summary

1,2,3,4-Tetramethoxy-10-methylacridone has been prepared and shown to be identical with a natural specimen of melicopicine. The four related isomeric trimethoxy-10-methylacridones have been prepared and the 2,3,4-trimethoxy derivative was shown to be identical with the trimethoxy derivative obtained from evoxanthine. A new synthesis of 10-methylacridones is described.

## I. INTRODUCTION

Melicopicine (I) was shown by Crow and Price(1) to be 1,2,3,4-tetramethoxy-10-methylacridone and as both melicopine and melicopidine have been converted to melicopicine a synthesis of this would confirm the structures of these three major alkaloidal constituents of *Melicope fareana* F. v. M. Also evoxanthine (II) isolated from *Evodia xanthoxyloides* F. v. M. was shown by Hughes and Neill(2) to be 2,3-methylenedioxy-4-methoxy-10-methylacridone, from which has been obtained 2,3,4-trimethoxy-10-methylacridone. Since no homonuclear trimethoxy-10-methylacridones have been described, the preparation of the four possible isomers has been carried out.



Preliminary experiments in the preparation of the diphenylamine acids by condensation of *o*-halogenobenzoic acids with methoxylated amines yielded intractable tars, but condensation of the methoxylated iodo- or bromobenzene with anthranilic acid in boiling nitrobenzene(3) went smoothly giving fair yields of diphenylamine acids.

Of the required iodo compounds the 2,3,5-trimethoxyiodobenzene, 2,4,5-trimethoxyiodobenzene and 2,3,4,5-tetramethoxyiodobenzene had not been previously described. On iodination with iodine and mercuric oxide(4), 1,2,4-trimethoxybenzene contrary to the finding of Erdtman(5) readily yielded 2,4,5-trimethoxyiodobenzene. *v*-Tetramethoxybenzene has only been dibrominated or dinitrated :

\* Department of Organic Chemistry, University of Sydney.

but as it is usually more difficult to introduce a second iodine atom into a phenolic ether, tetramethoxybenzene was iodinated with iodine and hydrogen peroxide(6) and yielded the monoiodinated product.

2,3,5-Trimethoxyiodobenzene cannot be prepared by direct iodination of readily available materials. 5-Iodovanillin which on treatment with alkali and hydrogen peroxide by Dakin's(7) method was expected to yield the requisite quinol failed to react (cf. Dakin 8). 5-Bromovanillin, however, gave a 45 per cent. overall yield of 2,3,5-trimethoxybromobenzene by a modification of this method.

Of the various methods of preparing 10-methylacridones that of Albert(9) via the route 5-aminoacridine  $\rightarrow$  5-aminoacridinium methiodide  $\rightarrow$  10-methylacridone appeared to be the most convenient, but this method failed with the tri- and tetramethoxy derivatives yielding only the acridone on hydrolysis with sodium carbonate.

Gleu and Nitzsche(10) prepared the monomethoxy-10-methylacridones by the action of dimethyl sulphate on the potassium salt of the acridone (cf. Drummond and Lahey(11) for preparation of 2,4-dimethoxy-10-methylacridone); but repetition of their work in the case of the 1-methoxy derivative gave only the unchanged acridone (Matheson, personal communication). In view of the failure of these methods another route to 10-methylacridones was sought. Lehmstedt(12) prepared 10-methylacridone by heating 5-methoxyacridone in a sealed tube at 200 °C. Gleu and Nitzsche(10) mention that though this reaction should proceed more easily at lower temperatures in the presence of methyl iodide it is of little importance compared with the preparation of 10-methylacridones by methylation of the potassium salt of the acridone. Neither of these authors appears to have attempted this transformation in the presence of methyl iodide. When 5-methoxyacridine was heated in a sealed tube with methyl iodide at 100 °C. it isomerized to 10-methylacridone in good yield.

This method was then applied to the preparation of the substituted 10-methylacridones and gave good yields in each case. 1,2,3,4-Tetramethoxy-10-methylacridone was shown to be identical with melicopicine by mixed melting point with a natural specimen and also by demethylation with alcoholic hydrochloric acid to 1,2,3-trimethoxy-4-hydroxy-10-methylacridone which was identified by melting point and mixed melting point with an authentic specimen of *normelicopicine*.

In a like manner 2,3,4-trimethoxy-10-methylacridone was shown to be identical with the trimethoxy derivative obtained from evoxanthine and the demethylated derivatives of both were also identical.

## II. EXPERIMENTAL

All melting points are uncorrected.

### (i) 10-Methylacridone

Anhydrous 5-methoxyacridine (0.5 g.)(13) was heated in a sealed tube with methyl iodide (1.5 ml.) in a boiling water-bath for 3 hours and the methyl iodide evaporated. The product crystallized from alcohol in long yellow needles. Its solution in alcohol had an intense violet fluorescence. Its m.p. and mixed m.p. with an authentic specimen was 203–204 °C.

(ii) *2,3,4,5-Tetramethoxyiodobenzene*

To tetramethoxybenzene (17.7 g.) (7), iodine (11.4 g.; 0.5 mol.) and concentrated sulphuric acid (1.5 ml.) in alcohol (150 ml.) at about 50 °C. was added 30% hydrogen peroxide (20 ml.) in small portions during 20 minutes with constant shaking, followed by heating on the water-bath for 10 minutes. The mixture was cooled, diluted with water (500 ml.), extracted with ether and the ether extract washed with sodium thiosulphate and water. The iodo compound distilled as a red-orange slightly viscous heavy liquid, b.p. 141–143 °C./0.35 mm. Yield 23.9 g. (82.5%),  $n_D^{23}$  1.5633.

Found: I, 32.0%.

Calculated for  $C_{10}H_{13}O_4I$ : I, 32.4%.

(iii) *2,3,4,5-Tetramethoxydiphenylamine-2'-Carboxylic Acid*

Tetramethoxyiodobenzene (4.3 g.), anthranilic acid (2 g.), anhydrous potassium carbonate (2 g.), and copper powder (0.1 g.) were refluxed in nitrobenzene (12 ml.) for 3 hours, the nitrobenzene removed by steam distillation, the residual solution filtered, and made acid to congo paper with hydrochloric acid. The product after reprecipitation from sodium bicarbonate was dried and crystallized from benzene-light petroleum as pale brown needles, m.p. 159–160 °C. Yield 1.9 g. (42%).

Found: N, 4.5;  $CH_3O$ , 36.5%.

Calculated for  $C_{17}H_{19}O_6N$ : N, 4.2;  $CH_3O$ , 37.2%.

(iv) *1,2,3,4-Tetramethoxy-5-Aminoacridine*

Tetramethoxydiphenylaminocarboxylic acid (3 g.) was converted to the 5-aminoacridine by the method of Albert and Ritchie (14). The product crystallized from light petroleum in bright yellow needles, m.p. 144–146 °C.

Found: N, 9.0%.

Calculated for  $C_{17}H_{18}O_4N_2$ : N, 8.9%.

This readily formed the acridinium methiodide which on hydrolysis with dilute sodium carbonate gave 1,2,3,4-tetramethoxyacridone, m.p. and mixed m.p. with product from (vi) 202–203 °C.

(v) *1,2,3,4-Tetramethoxy-10-Methylacridone (Melicopine)*

Tetramethoxydiphenylaminocarboxylic acid (2 g.) was converted to the 5-chloroacridine by refluxing with phosphorus oxychloride (80 ml.) for 3 hours in an oil-bath at 135–145 °C., the excess phosphorus oxychloride removed under vacuum, the residue taken up in chloroform, and made alkaline with ice cold ammonia. The chloroform was removed and the solid chloroacridine dissolved in dry methanol and added to a solution of sodium methoxide (from 0.4 g. sodium) in dry methanol and refluxed on a water-bath for 2 hours. The residue after distillation of the methanol was diluted with water and extracted with chloroform. Evaporation of the chloroform gave the crude 1,2,3,4,5-pentamethoxyacridine as a dark oil, which could not be readily purified. It was heated with methyl iodide (2–3 ml.) in a sealed tube for 4 hours. The methyl iodide was evaporated, the dark residue dissolved in benzene, and chromatographed on alumina. From the eluate on concentration and addition of light petroleum, crystallized the 1,2,3,4-tetramethoxy-10-methylacridone in large square plates having m.p. 131–133 °C. either alone or admixed with an authentic specimen of melicopine from *M. fareana*. Yield 0.3 g. (15%).

Found: O, 23.8;  $CH_3O$ , 37.1;  $CH_3N$ , 7.2%.

Calculated for  $C_{18}H_{19}O_5N$ : O, 24.2;  $CH_3O$ , 37.6;  $CH_3N$ , 9.0%.

The identity of the synthetic material was confirmed by demethylation with alcoholic hydrochloric acid for 1 hour to normelicopine which crystallized from aqueous alcohol in long orange needles, m.p. 125–127 °C. either alone or admixed with an authentic specimen.

(vi) *1,2,3,4-Tetramethoxyacridone*

1,2,3,4-Tetramethoxy-5-chloroacridine prepared as in (v) was refluxed for half an hour with concentrated hydrochloric acid (30 ml.), cooled, and made alkaline with sodium hydroxide. The



resinous precipitate was extracted with 10% hydrochloric acid leaving a crystalline residue which crystallized in large glistening golden plates from aqueous alcohol, m.p. 202–203 °C.

Found: C, 64.3; H, 5.4; N, 4.7%.

Calculated for  $C_{17}H_{17}O_5N$ : C, 64.8; H, 5.4; N, 4.4%.

(vii) *5-Nitro-1,2,3-Trimethoxybenzene*

There is no satisfactory preparation of this compound recorded in the literature. To a stirred mixture of concentrated nitric acid (100 ml.) and glacial acetic acid (200 ml.) cooled in ice, was added in small portions trimethylgallic acid (50 g.)(15). The solution turned dark red and towards the end of the addition carbon dioxide was evolved. The mixture was then stirred for half an hour at room temperature and poured into water (1.5 l.). The product was washed with sodium carbonate solution and was sufficiently pure for the reduction, m.p. 95–97 °C. Yield 31.8 g. (64%).

(viii) *3,4,5-Trimethoxyaniline*

This has been prepared by reduction with stannous chloride(16) but is more conveniently prepared as follows:

The 5-nitro-1,2,3-trimethoxybenzene (10 g.) was refluxed with water (50 ml.), iron powder (25 g.), and ferrous sulphate (2 g.) for 4 hours. The product was diluted with boiling water (50 ml.) and filtered through a preheated funnel. The amine crystallized on cooling, m.p. 114 °C. Yield 7.1 g. (83%).

(ix) *3,4,5-Trimethoxydiphenylamine-2'-Carboxylic Acid*

3,4,5-Trimethoxyiodobenzene (15 g.)(17) was condensed with anthranilic acid as in (iii) and the acid obtained was crystallized with difficulty from benzene-light petroleum in off white needles, m.p. 164–166 °C. Yield 3.5 g. (23%).

Found: N, 4.9;  $CH_3O$ , 31.3%.

Calculated for  $C_{16}H_{17}O_5N$ : N, 4.6;  $CH_3O$ , 30.7%.

(x) *2,3,4-Trimethoxy-10-Methylacridone*

This was prepared as in (v) from 3,4,5-trimethoxydiphenylamine-2'-carboxylic acid. After being chromatographed on alumina, it was crystallized from benzene-light petroleum in large yellow plates, m.p. either alone or admixed with a specimen prepared from evoxanthine 168.5–170 °C.

Found: C, 67.9; H, 5.8; N, 4.7;  $CH_3O$ , 30.5;  $CH_3N$ , 8.2%.

Calculated for  $C_{17}H_{17}O_4N$ : C, 68.2; H, 5.7; N, 4.7;  $CH_3O$ , 31.1;  $CH_3N$ , 9.7%.

The identity was confirmed by its demethylation on heating with alcoholic hydrochloric acid for  $\frac{3}{4}$  hour to 2,3-dimethoxy-4-hydroxy-10-methylacridone which crystallized from alcohol in long golden needles, m.p. 176–177 °C. either alone or admixed with a specimen prepared from evoxanthine.

(xi) *2,3,4-Trimethoxyacridone*

This was prepared as in (vi) and crystallized from aqueous alcohol or benzene-light petroleum as pale cream needles, m.p. 208–209 °C.

Found: C, 68.1; H, 5.7; N, 4.7;  $(CH_3O)$ , 31.5%.

Calculated for  $C_{16}H_{15}O_4N$ : C, 67.4; H, 5.3; N, 4.9;  $CH_3O$ , 32.7%.

(xii) *2,4,5-Trimethoxyiodobenzene*

To 1,2,4-trimethoxybenzene (5 g.) and yellow mercuric oxide (4.9 g.; 0.75 mol.) in alcohol (20 ml.), was added iodine (7.6 g.; 1 mol.) in three portions, with shaking, the temperature being kept at 45–55 °C. The mixture was filtered, cooled, diluted with water, and sodium bisulphite

added. The iodo compound crystallized from aqueous alcohol or light petroleum as colourless prisms, m.p. 71–72 °C. Yield 6.2 g. (36%).

Found : I, 43.7%.

Calculated for  $C_9H_{11}O_3I$  : I, 43.2%.

(xiii) *2,4,5-Trimethoxydiphenylamine-2'-Carboxylic Acid*

2,4,5-Trimethoxyiodobenzene (10 g.) was condensed with anthranilic acid as in (iii) and the product crystallized from toluene as light brown needles, m.p. 183–183.5 °C. (sinters 170 °C.) Yield 4.5 g. (41%).

Found : N, 4.7 ;  $CH_3O$ , 30.0%.

Calculated for  $C_{16}H_{17}O_5N$  : N, 4.6 ;  $CH_3O$ , 30.7%.

(xiv) *1,3,4-Trimethoxy-10-Methylacridone*

2,4,5-Trimethoxydiphenylamine-2'-carboxylic acid was treated as in (v) and the product after being chromatographed on alumina crystallized from benzene-light petroleum in bright yellow needles, m.p. 113–115 °C. It had an intense yellow-green fluorescence in benzene and yellow in alcohol.

Found : C, 67.9 ; H, 5.7 ; N, 4.8 ;  $CH_3O$ , 30.1 ;  $CH_3N$ , 9.7%.

Calculated for  $C_{17}H_{17}O_4N$  : C, 68.2 ; H, 5.7 ; N, 4.7 ;  $CH_3O$ , 31.1 ;  $CH_3N$ , 9.7%.

(xv) *1,3,4-Trimethoxyacridone*

This was prepared as in (vi) from 2,4,5-trimethoxydiphenylamine-2'-carboxylic acid and crystallized from water in fine yellow needles, m.p. 162–163 °C.

Found : C, 67.2 ; H, 5.2 ; N, 5.2 ;  $CH_3O$ , 32.2%.

Calculated for  $C_{16}H_{15}O_4N$  : C, 67.4 ; H, 5.3 ; N, 4.9 ;  $CH_3O$ , 32.7%.

(xvi) *1,3,4-Trimethoxy-5-Aminoacridine*

2,4,5-Trimethoxydiphenylamine-2'-carboxylic acid was converted as in (iv) to the 5-aminoacridine which crystallized from benzene-light petroleum as dark red prisms, m.p. 185–187 °C.

Found : N, 9.9%.

Calculated for  $C_{16}H_{16}O_3N$  : N, 9.9%.

This readily formed a methiodide which on refluxing with dilute sodium carbonate gave the 1,3,4-trimethoxyacridone identified by m.p. and mixed m.p. with that obtained in (xv).

(xvii) *2,3,5-Trimethoxybromobenzene*

To a solution of 5-bromovanillin (53 g.) in N/10 caustic potash (480 ml.) stirred in an atmosphere of coal gas was added 30% hydrogen peroxide (49 ml. ; 1 mol.) in one portion followed by more 6% hydrogen peroxide (245 ml.) added dropwise during an hour. The mixture was then stirred for a further 2 hours when some solid separated, then saturated with sulphur dioxide, cooled, and ether extracted. The dark solid obtained on evaporation was dissolved in excess sodium hydroxide (10%) containing sodium hydrosulphite (4–5 g.) and excess dimethyl sulphate added. The oil was ether extracted and distilled under reduced pressure as a colourless oil, b.p. 106–108 °C./0.3 mm., 170–173 °C./25 mm. Yield 25.5 g. (45%).

Found : Br, 48.3%.

Calculated for  $C_8H_9Br$  : Br, 48.8%.

(xviii) *2,3,5-Trimethoxydiphenylamine-2'-Carboxylic Acid*

2,3,5-Trimethoxybromobenzene (10 g.) was treated as in (iii) and the product crystallized from benzene in pale fawn cubes, m.p. 166–167 °C. Yield 8.5 g. (70%).

Found : N, 4.7 ;  $CH_3O$ , 30.6%.

Calculated for  $C_{16}H_{17}O_5N$  : N, 4.6 ;  $CH_3O$ , 30.7%.

(xix) *1,2,4-Trimethoxy-10-Methylacridone*

2,3,5-Trimethoxydiphenylamine-2'-carboxylic acid (2 g.) was treated as in (v) and the 1,2,4-trimethoxy-10-methylacridone after being chromatographed on alumina crystallized from benzene-light petroleum in long golden needles, m.p. 141–142 °C. The alcohol solution was yellow with a yellow-green fluorescence.

Found: C, 67.6; H, 5.8; N, 5.0; CH<sub>3</sub>O, 31.1; CH<sub>3</sub>N, 6.5%.

Calculated for C<sub>17</sub>H<sub>17</sub>O<sub>4</sub>N: C, 68.2; H, 5.7; N, 4.7; CH<sub>3</sub>O, 31.1; CH<sub>3</sub>N, 9.7%.

(xx) *1,2,4-Trimethoxyacridone*

This was prepared as in (vi) from 2,3,5-trimethoxydiphenylamine-2'-carboxylic acid (5 g.). It crystallized as long yellow needles from aqueous alcohol, m.p. 221–223 °C. Yield 3.7 g. (79%).

Found: C, 67.1; H, 5.3; N, 4.9; CH<sub>3</sub>O, 33.1%.

Calculated for C<sub>16</sub>H<sub>15</sub>O<sub>4</sub>N: C, 67.4; H, 5.3; N, 4.9; CH<sub>3</sub>O, 32.7%.

(xxi) *2,3,4-Trimethoxydiphenylamine-2'-Carboxylic Acid*

3,4,5-Trimethoxyiodobenzene (9 g.) prepared as in (ii) from pyrogallol trimethyl ether was converted to 2,3,4-trimethoxydiphenylamine-2'-carboxylic acid as in (iii) and the product crystallized from benzene-light petroleum as pale fawn needles, m.p. 160–161 °C. Yield 2.4 g. (27%).

Found: N, 4.8; CH<sub>3</sub>O, 30.9%.

Calculated for C<sub>16</sub>H<sub>17</sub>O<sub>5</sub>N: N, 4.6; CH<sub>3</sub>O, 30.7%.

(xxii) *1,2,3-Trimethoxy-10-Methylacridone*

As in (v) the 2,3,4-trimethoxydiphenylamine-2'-carboxylic acid (2 g.) was converted to 1,2,3-trimethoxy-10-methylacridone which after chromatographing on alumina crystallized in yellow rods from benzene-light petroleum, m.p. 116–118 °C.

Found: C, 67.9; N, 5.8; CH<sub>3</sub>O, 30.6; CH<sub>3</sub>N, 9.0%.

Calculated for C<sub>17</sub>H<sub>17</sub>O<sub>4</sub>N: C, 68.2; H, 5.7; N, 4.7; CH<sub>3</sub>O, 31.1; CH<sub>3</sub>N, 9.7%.

(xxiii) *1,2,3-Trimethoxyacridone*

Prepared as in (vi) from 2,3,4-trimethoxydiphenylamine-2'-carboxylic acid. The acridone crystallized from aqueous alcohol in cream needles or benzene-light petroleum as yellow needles m.p. 195–197 °C.

Found: C, 66.7; H, 5.4; N, 4.8; CH<sub>3</sub>O, 31.3%.

Calculated for C<sub>16</sub>H<sub>15</sub>O<sub>4</sub>N: C, 67.4; H, 5.3; N, 4.9; CH<sub>3</sub>O, 32.7%.

## IV. ACKNOWLEDGMENTS

The authors are indebted to Miss J. Fildes and Miss B. Naylor, micro-analysts, University of Sydney, for the analyses and to Dr. J. R. Price, C.S.I.R.O., for the sample of melicopicine.

## V. REFERENCES

- (1) CROW, W. D., and PRICE, J. R.—*Aust. J. Sci. Res. A* **2**: 282 (1949).
- (2) HUGHES, G. K., and NEILL, K. G.—*Aust. J. Sci. Res. A* **2**: 429 (1949).
- (3) GOLDBERG, I.—*Ber. dtsh. chem. Ges.* **39**: 1691 (1906).
- (4) JURD, L.—*Aust. J. Sci. Res. A* **2**: 246 (1949).
- (5) ERDTMAN, H.—*Proc. Roy. Soc. A* **143**: 215 (1933-4).
- (6) JURD, L.—*Aust. J. Sci. Res. A* **2**: 247 (1949).
- (7) BAKER, W., JUKES, E. H. T., and SUBRAHMANYAN, C. A.—*J. Chem. Soc.* **1934**: 1681 (1934).

- (8) DAKIN, H. D.—*Amer. Chem. J.* **42**: 493 (1909).
- (9) ALBERT, A., and RITCHIE, B.—*J. Chem. Soc.* **1943**: 458 (1943).
- (10) GLEU, K., and NITZSCHE, Z.—*J. prakt. Chem.* **153**: 200 (1939).
- (11) DRUMMOND, L. J., and LAHEY, N.—*Aust. J. Sci. Res. A* **2** (1949).
- (12) LEHMSTEDT, K.—*Ber. dtsh. chem. Ges.* **68**: \*1464 (1935).
- (13) DROSDOV, N. J., and CHERNSTOV, O. M.—*J. Gen. Chem. U.S.S.R.* **14**: 181 (1944). (*Chem. Abstr.* **39**: 2290 (1945).)
- (14) ALBERT, A., and RITCHIE, B.—*Organic Synth.* **22**: 5 (1942).
- (15) MAUTHNER, F.—*Org. Synth.* **6**: 96 (1926).
- (16) WILL, C.—*Ber. dtsh. chem. Ges.* **21**: 613 (1888).
- (17) GRAEBE, C., and SUTER, M.—*Liebigs Ann.* **340**: 230 (1905).



# THE EXISTENCE OF A URONIC ACID ESTER IN YOUNG WOOD OF *EUCALYPTUS REGNANS*

By D. H. FOSTER,\* GERDA SCHWERIN,\* and W. E. COHEN\*

[Manuscript received April 7, 1950]

## Summary

A comparison has been made between the hydrolysis of the hemicellulose in young *Eucalyptus regnans* F. v. M. and the hemicellulose in holocellulose prepared from the same wood sample. The rates of hydrolysis of the pentosan and uronic acid constituents have been compared. The results have been interpreted in terms of an ester in which uronic acid is thought to be involved.

## I. INTRODUCTION

The exceedingly complex arrangement of wood polysaccharides is further complicated by the presence of lignin which exerts a pronounced influence on their chemical reactions. Preliminary experiments have indicated that hydrolysis of uronic acid in wood lagged behind the hydrolysis of what was considered to be the associated pentosan. It was thought that if lignin inhibited the liberation of uronic acid during dilute acid hydrolysis, then the phenomenon could best be studied by comparing the rates of hydrolysis of uronic acid anhydride by reference to the rates of pentosan hydrolysis in a wood sample and in holocellulose prepared from the same sample, i.e. in wood and in delignified wood. It is important that pentosan hydrolysis should be used as a reference for uronic acid anhydride hydrolysis in each case. If the absolute rates of uronic acid anhydride hydrolysis were compared, there would be some uncertainty as to what extent the result would be affected by the physical differences in the raw materials. Such differences might, for example, alter the rate of diffusion of the hydrolysing acid.

## II. MATERIALS USED

Wood from *Eucalyptus regnans* trees, about nine years old, was used. Within a few days of felling, the pith and outer sapwood were removed, the residue converted to sawdust and immediately extracted with cold water, cold alcohol, and again cold water, successively over a period of several days in order to remove tannin, kino, and other extractives. The final water wash served only to remove alcohol from the wood. In order to avoid chemical modification of the wood, no hot solvents were used.

The extracted sawdust was air-dried, comminuted in the Wiley mill, screened, and the 60–80 mesh fraction accepted for investigation.

A portion of this material was used to prepare holocellulose in bulk by the method of Wise, Murphy, and d'Addieco(1). Wood (118.5 g., oven-dry) was

\* Division of Forest Products, C.S.I.R.O., Melbourne.

given five treatments, each with 35 g. sodium chlorite and 6 ml. glacial acetic acid, to yield 80.45 per cent. of holocellulose. The original wood contained 23.4 per cent. lignin.

### III. EXPERIMENTAL PROCEDURE AND RESULTS

Hydrolysis was carried out, at 91–92 °C., in 500 ml. Quickfit Erlenmeyer flasks equipped with air condensers.

TABLE 1

ANALYSIS OF WOOD TREATED FOR VARIOUS TIME PERIODS WITH N SULPHURIC ACID AT 91–92 °C.  
All results based on original extracted wood

Time (hr.)	Yield (%)	Pentosan Corrected for Furfural from Uronic Acid (%)	Average (%)	Uronic Acid (%)	Average (%)
0	100	17.41 17.69	17.55	3.48 3.56 3.50 3.72	3.56
0.33	96.3	16.65 16.35	16.50	3.02 3.58	3.30
0.50	94.45	15.86 15.95	15.91	3.38 3.36	3.37
0.72	92.0	14.38 14.49	14.44	3.16 3.14	3.15
1.0	89.92	13.19 13.33	13.26	3.40 3.21	3.30
2.0	83.55	9.01 8.88	8.95	2.72 3.06	
	83.9			2.50 2.70	2.76
	83.4			2.92 2.64	
3.0	79.43	6.28 6.09	6.19	2.12 2.16	2.07
	77.78			1.87 1.93	
4.0	76.68	5.37 5.62	5.50	1.78 1.68	1.73

Wood (12 g., equal to 10.98 g. oven-dry) or holocellulose (12 g., equal to 10.83 g. oven-dry) was transferred to the flask which was placed in an oil-bath,

and to which was then added N sulphuric acid (250 ml.) which had been preheated to approximately 95 °C. The hydrolysis reaction was terminated by adding cold distilled water (250 ml.) to the flask and filtering off the residue on a sintered glass funnel. After it had been washed with 1500 ml. of cold water during a period of about 1 hour, the residue was air-dried and weighed. Pentosan was estimated by a method developed in this Laboratory which is similar to that described by Dorée(2) except that thiobarbituric acid was used instead of phloroglucinol(3). Uronic acid anhydride was estimated by the method of Whistler, Martin, and Harris(4).

The results of these analyses, all based on 100 g. (oven-dry) extracted original wood, are given in Tables 1 and 2.

TABLE 2

ANALYSIS OF HOLOCELLULOSE TREATED FOR VARIOUS TIME PERIODS WITH N SULPHURIC ACID AT 91-92 °C.

All results based on original extracted wood

Time (hr.)	Yield (%)	Pentosan Corrected for Furfural from Uronic Acid (%)	Average (%)	Uronic Acid (%)	Average (%)
0	80.45	16.60 15.97	16.29	4.08 4.20	4.14
0.33	66.1	11.62 11.94	11.78	2.91 3.12	3.01
0.66	61.95	8.46 8.54	8.50	2.36 2.46	2.41
1.0	56.50	6.58 6.43	6.51	1.18 1.15	1.16
2.0	54.10	4.91 5.05	4.98	0.72 0.75	0.73
3.0	52.80 52.57	4.16 4.44	4.30	0.43 0.47 0.61 0.54	0.51
4.0	51.68	3.85 3.76	3.81	0.61 0.66	0.63

The residue, its pentosan and its uronic acid anhydride contents, all expressed as percentages of the original wood, have been plotted against time as shown in Figures 1, 2, and 3 respectively. Pentosan results are corrected for furfural from uronic acid by using the values given by Norris and Resch(5).

## IV. DISCUSSION

The overall effect of lignin removal on the rate of hydrolysis of wood carbohydrates is illustrated in Figure 1. The more rapid hydrolysis of pentosan in the holocellulose probably arose from the fact that the hydrolysing acid could diffuse more easily into the holocellulose because there was no lignin present to exert a protecting influence. After 1 to 2 hours of holocellulose hydrolysis most of the hemicellulose was removed and only the residual resistant carbohydrate was being attacked. Hence, from Figure 1, the amount of easily

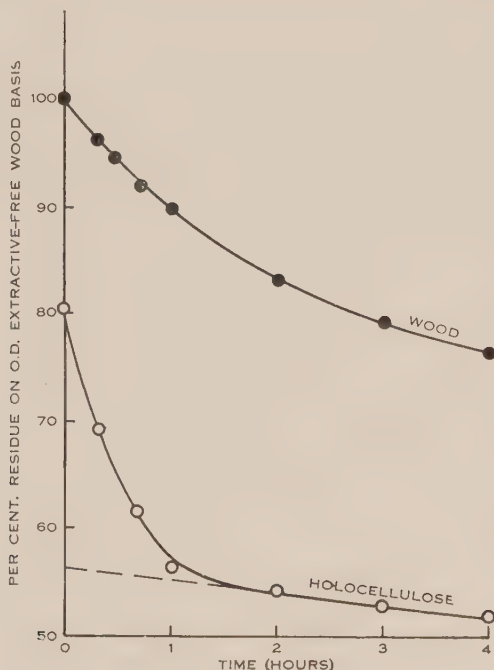


Fig. 1.—Total hydrolysis of wood and holocellulose with  $N$  sulphuric acid at 91–92 °C.

hydrolysable hemicellulose can be found by extrapolation to be approximately 23 ( $80.45 - 57.5$ ) per cent. of the original wood.

Reference to Figure 2 shows how pentosan is hydrolysed from both wood and holocellulose. A large proportion of the pentosan (11.4 per cent. on oven-dry wood basis) was hydrolysed from the wood in 3 hours, and approximately the same amount (9.8 per cent.) was hydrolysed from the holocellulose in one hour. The slightly lower figure for pentosan (9.8 as against 11.4 per cent.) hydrolysed from holocellulose is attributable to a pentosan loss incurred during preparation of the holocellulose. Reference to Figure 3 shows the relative rates of hydrolysis of uronic acid from wood and from holocellulose and it can be seen that uronic acid was hydrolysed from the wood considerably more slowly than it was from the holocellulose. There was also more uronic acid in the untreated holocellulose than there was in the original wood. This may



have been due to oxidation of primary hydroxyl groups in hexose units by the sodium chlorite used in the preparation of the holocellulose. The excess uronic acid in the holocellulose has not been considered in comparing the amounts hydrolysed from wood and holocellulose. Thus 1.45 (3.55—2.1) per cent.

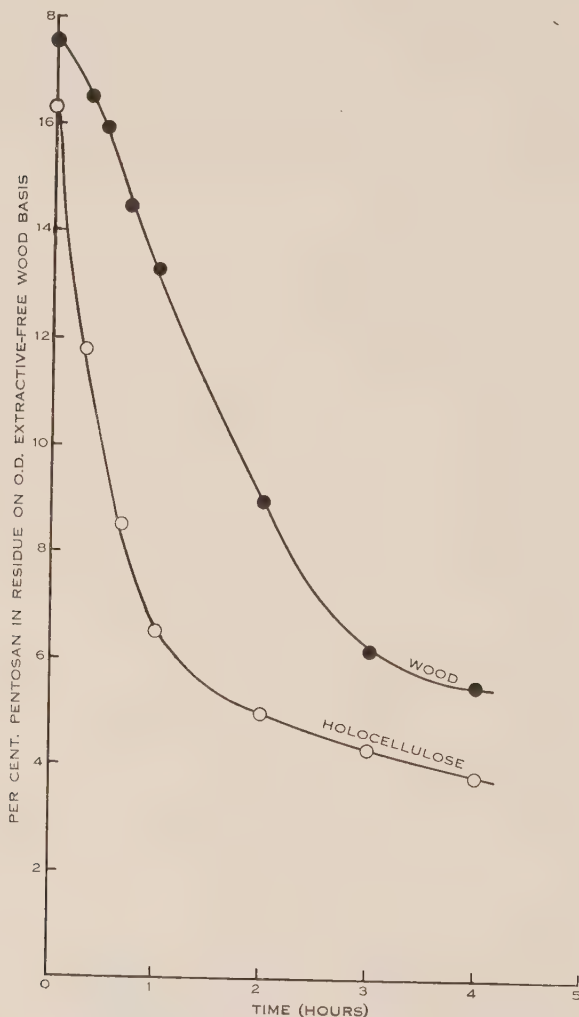


Fig. 2.—Pentosan hydrolysed from wood and holocellulose with N sulphuric acid at 91–92 °C.

uronic acid anhydride was hydrolysed from the wood in 3 hours and 2.35 (3.55—1.2) per cent. from the holocellulose in 1 hour.

Hence the molar ratios are :

For wood :

$$\frac{\text{uronic acid anhydride}}{\text{pentosan}} = \frac{1.45}{11.4} \times \frac{132}{176} = \frac{1}{10.5} = 0.095.$$

For holocellulose :

$$\frac{\text{uronic acid anhydride}}{\text{pentosan}} = \frac{2.35}{9.8} \times \frac{132}{176} = \frac{1}{5.6} = 0.180.$$

The comparatively slow hydrolysis of uronic acid in the wood may be accounted for by assuming that it is chemically bound to some portion of the wood through a link other than the glycosidic link by which it is held to xylan units. Physical protection of the uronic acid anhydride by lignin is unlikely

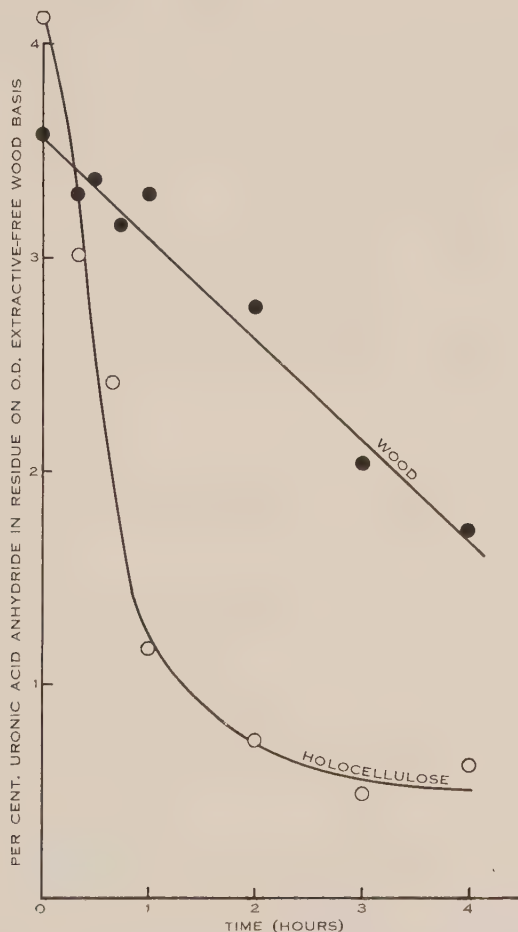


Fig. 3.—Uronic anhydride hydrolysed from wood and holocellulose with N sulphuric acid at 91–92 °C.

because the uronic acid anhydride and pentosan are known to be closely associated and it would be expected that protection would extend to both these constituents in the wood. An ether link formed through a secondary hydroxyl belonging to uronic acid and another wood constituent is unlikely because such a link should be very resistant to dilute acid whereas the link under discussion is slowly broken by the dilute sulphuric acid used in the experiment. However,

it is possible that there is an ester link between the uronic acid carboxyl and a hydroxyl belonging to either lignin or carbohydrate or to both.

Sarkar *et al.*(8) considered that they had obtained evidence of an ester link between uronic acid and lignin in jute fibre. They defatted jute fibre and delignified it with sodium chlorite and acetic acid (pH 3-5) on a boiling water-bath. After delignification it was found that the acid value as determined by the silver *o*-nitrophenate method(7, 8) almost doubled. The acid value was also doubled when the fibre was given a treatment with 0.25 per cent. sodium hydroxide for 1 hour at ordinary temperature and it remained almost unchanged on delignification. Furthermore, the acid value of defatted jute did not increase on boiling the fibre with 0.1N mineral acid for half an hour under reflux. Thus it appears that a mild hydrolysis was insufficient to break the supposed ester link in jute fibre. This agrees with the finding reported here, viz. that uronic acid is only slowly liberated during the hydrolysis of *E. regnans* wood with N sulphuric acid at 91-92 °C.

It is not considered that Sarkar's results preclude the possibility of a carbohydrate-carbohydrate ester link, because treatment with sodium chlorite may involve reactions other than those associated with delignification. However, the supposition of an ester link between uronic acid and lignin in *E. regnans* wood conforms with the following:

- (a) A large proportion of the hemicelluloses may be extracted from wood with dilute alkali.
- (b) Gray *et al.*(10), Eastham *et al.*(11), and Freudenberg(12) consider that lignin contains primary and secondary hydroxyls: an ester linkage with lignin could involve these.
- (c) Harris, Sherrard, and Mitchell(13) found that an acid hydrolysis was necessary in order further to methylate partially methylated maple wood and concluded that lignin must be attached to some of the carbohydrates, possibly hemicelluloses, in woods.
- (d) Dilute acid hydrolysis renders a large portion of hardwood lignin soluble in hot methanol: this has been found to be true with a eucalypt wood(14).

During the preparation of holocellulose by the method of Van Beckum and Ritter an increase in nitrogen content occurs. Reynolds(15) found that chlorination methods of holocellulose preparation gave a product containing about 0.5 per cent. nitrogen. In investigating the reason for this nitrogen retention Reynolds found that little or no nitrogen was introduced into *E. regnans* wood by the action of ammonia or ethanolamine unless the treatment was preceded by chlorination. Furthermore, Wise, Murphy, and d'Addieco(1) found that holocellulose prepared with sodium chlorite had a low nitrogen content but readily took up 0.3 per cent. nitrogen on being treated with alcoholic ethanolamine. A possible explanation of these reactions is that the ester linkage in which the uronic acid takes part is broken and the uronic acid forms a lactone which then forms an amide.

Brauns and Yirak(16) consider a glucosidic or ether type of lignin-carbohydrate linkage to be likely and an ester linkage to be unlikely. Fisher, Hawkins, and Hibbert(17) have also suggested a glucosidic linkage, but such a linkage would not be broken easily by alkali, and an ether linkage would be difficult to break with dilute acid or with alkali. However, it is not denied that there may be some ether or glucosidic linkages because there is some lignin in wood which is difficult to render soluble in methanol by acid treatment, and which is hard to remove in alkaline pulping. Such linkages may involve the resistant cellulose.

## V. CONCLUSION

The experiments described above indicate that the carboxyl of uronic acid is combined in ester formation with carbohydrate or with lignin and, in view of the results which have been obtained by other workers, it seems that the combination with lignin is very likely, but that the carbohydrate hydroxyl-uronic acid ester might easily co-exist.

## VI. REFERENCES

- (1) WISE, L. E., MURPHY, M., and D'ADDIECO, A. A.—*Paper Tr. J.* **122** (2) : 35-43 (1946).
- (2) DORÉE, D. C.—“The Methods of Cellulose Chemistry.” p. 365. (Chapman and Hall Ltd. : London, 1933.)
- (3) MACKNEY, A. W., and REYNOLDS, T. M.—*J. Coun. Sci. Industr. Res. Aust.* **11** : 333-48 (1938).
- (4) WHISTLER, R. C., MARTIN, A. R., and HARRIS, M.—*J. Res. Nat. Bur. Stand.* **24** : 13-23 (1940).
- (5) NORRIS, F. W., and RESCH, C. E.—*Biochem. J.* **29** : 1590-96 (1935).
- (6) SARKAR, P. B., CHATTERGEE, H., MAZUMDAR, A. K., and PAL, B.—*J. Text. Inst.* **39** (1) : T.1-7 (1948).
- (7) SOOKNE, A. M., and HARRIS, M.—*J. Res. Nat. Bur. Stand.* **26** : 205 (1941).
- (8) SARKAR, P. B., CHATTERGEE, H., and MAZUMDAR, A. K.—*J. Text. Inst.* **38** (9) : T.318-32 (1947).
- (9) PULP and PAPER RES. INST. CANAD.—7th Ann. Rep., p. 28, Montreal (1946-47).
- (10) GRAY, K. R., KING, E. G., BRAUNS, F., and HIBBERT, H.—*Canad. J. Res. B* **13** : 35-47 (1935).
- (11) EASTHAM, A. M., FISHER, H. E., KULKA, M., and HIBBERT, H.—*J. Amer. Chem. Soc.* **66** (1) : 26-32 (1944).
- (12) FREUDENBERG, F.—*Das Papier* **1** (11/12) : 209 (1947).
- (13) HARRIS, E. E., SHERRARD, E. C., and MITCHELL, R. L.—*J. Amer. Chem. Soc.* **56** : 889-93 (1934).
- (14) RALPH, B. J., and WARDROP, A. B.—*J. Aust. Chem. Inst.* **13** : 144-55 (1946).
- (15) REYNOLDS, T. M.—Nitrogen content of holocellulose fractions prepared by various methods. Coun. Sci. Industr. Res. Aust. Div. Forest Products : Sub-Proj. C.1-5, Prog. Rep. No. 5 (1940). (Mimeo.)
- (16) BRAUNS, F. E., and YIRAK, J. J.—*Paper Tr. J.* **125**, 12 : 55-60 (1947).
- (17) FISHER, J. H., HAWKINS, W. C., and HIBBERT, H.—*J. Amer. Chem. Soc.* **63** (11) : 3031-5 (1941).



## SHORT COMMUNICATIONS

### NUCLEAR EMULSION TECHNIQUE\*

By V. D. HOPPER†

Experiments on cosmic rays are being carried out using Ilford G.5 Nuclear Research emulsions, which are capable of recording particles at very low ionizing power. With careful development it is possible to record tracks of particles moving at velocities giving minimum ionization although it is still impossible to trace slow electron paths with certainty. In carrying out experiments with these emulsions in Australia, it is necessary to transport the plates from England by plane or ship and during transit the plates are affected by cosmic radiation and tropical conditions. Methods at present used for the eradication of background suffer from the disadvantage that the sensitivity of the emulsion is reduced in the process.

The background in the plates due to cosmic rays depends on the age of the emulsion and until the emulsions can be prepared in Australia it is preferable to transport them by plane. In this way emulsions two or three weeks old can be used for experiments. If the plates are transported by ship, the age of the plates is six or more weeks on arrival and the background density when finally developed is therefore heavy, making it more difficult to distinguish tracks of minimum ionizing particles. On development of plates transported by plane it was found that a dense surface blackening had occurred which was difficult to remove if left until the fixation and drying processes were completed. It was found that this surface layer could be removed by wiping the surface of the plate during the development process and plates treated in this way showed minimum ionizing tracks through the emulsion except for a layer a few microns thick near the surface. An example from an emulsion treated in this way is shown in Plate 1. The mosaic shows a  $\pi$ -meson coming to rest, disintegrating into a  $\mu$ -meson which in turn comes to rest after a range of  $600\ \mu$  and decays into an electron. The electron produces a track at minimum ionization and has in this case a grain density of approximately 24 grains/100  $\mu$ .

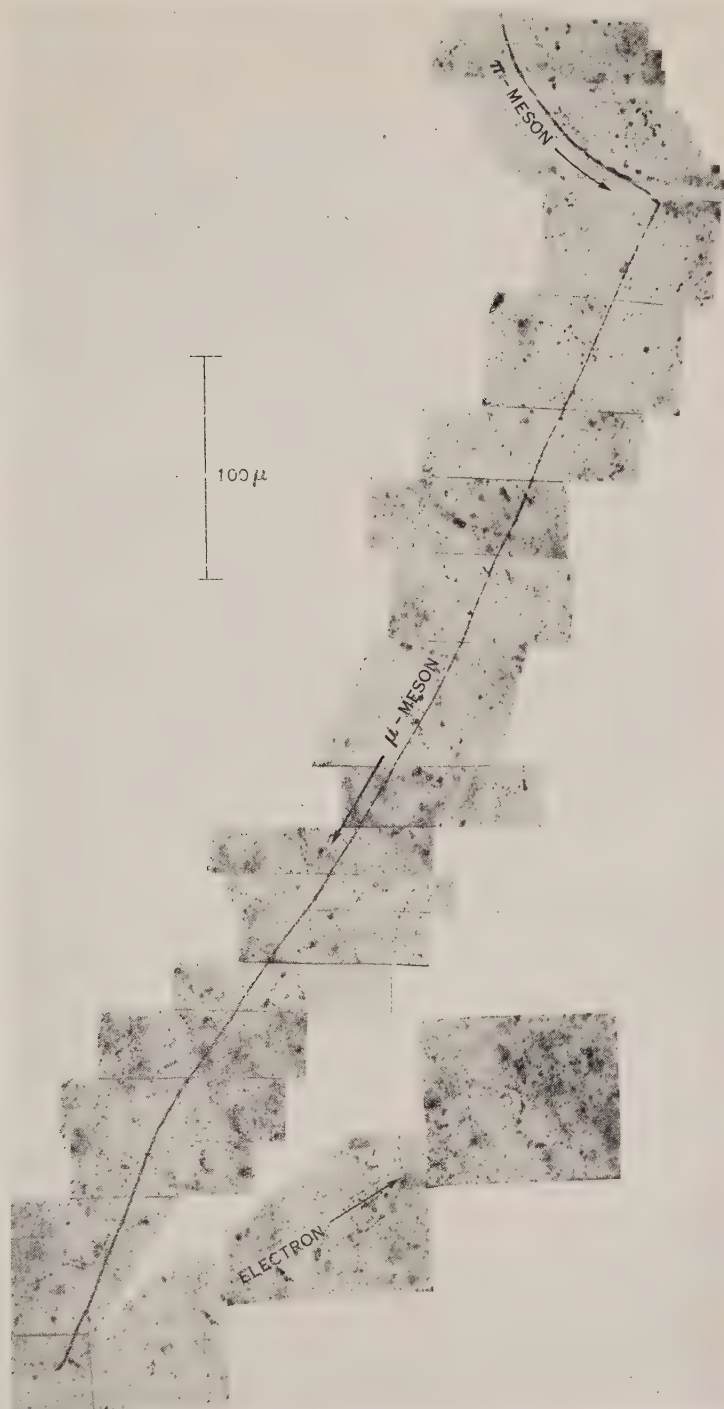
#### *Development of 400 $\mu$ Thick Emulsion*

For emulsions 400  $\mu$  thick on glass the following method was adopted.

Freshly made Agfa developer of 4 per cent. concentration was cooled below  $4^\circ\text{C}$ . by placing the developing dish in an ice water mixture(1) and the plate allowed to soak in this for 2 hours. It was then placed in fresh developer

\* Manuscript received July 6, 1950.

† Physics Department, University of Melbourne.





at 18 °C. and allowed to develop for another 50 minutes. No stirring was employed during the development. It was rinsed and bathed for 1 hour in a 2 per cent. acetic acid stop bath. The plate was then placed in a hypo solution (400 g. of hypo in 1 litre of water) and stirred with an electrically operated rocker. After 15 minutes the surface was wiped with cotton wool to remove the dense black surface layer which had accumulated during the development process. The fixer was changed after 15 minutes, 30 minutes, 1 hour, 1 hour, 1 hour, 2 hours, 4 hours, and the plate left in the fixer for another 12 hours. At this stage a light source could be seen through the plate but the emulsion had a milky appearance. The milky background could be reduced by leaving the plate in the fixer for a much longer period but it has been found that this does not completely remove it. During the fixation process the rate of solution of silver halide was studied by slicing off thin sections of emulsion near the edge and observing these with a microscope, the cut of the section being at right angles to the optic axis of the microscope. The undissolved silver appeared opaque and in the remaining volume tracks of ionizing particles were visible.

The plate was next washed in running water for 4 hours and the milkiness of the emulsion diminished slightly. It was found possible to reduce this considerably by the following process but there is additional risk of emulsion distortion. The plate was placed in acid fixer and the temperature of the solution gradually raised to 30 °C. In less than 30 minutes the milkiness had practically disappeared and the solution was allowed to cool to room temperature. (The major distortion of the emulsion occurs during the cold water washing process following the acid fixer treatment and is due to the additional swelling of the emulsion in water. By reducing the concentration of the acid fixer gradually this danger of distortion can be removed.) The plate was dried vertically (drying time approximately 48 hours) and after drying the surface was cleaned by wiping it with a soft chamois leather damped with benzine.

### Results

Table 1 gives a summary of data obtained from two plates. Plate No. 1 had been flown out by plane and had been kept at ground level whereas plate No. 2 had been flown out by plane but had been given an additional cosmic ray exposure on aircraft flights above 25,000 feet for approximately 7 hours.

These preliminary results indicate that a useful study can be made in Australia of cosmic rays at high altitudes using these plates. Exposures are

TABLE 1

Plate No.	Area (cm. <sup>2</sup> )	Average Prongs/Star	Age of Plate (days)	Number of Stars with following Number of Prongs :									
				3	4	5	6	7	8	9	10	11	12
1	30	3.83	42	10	6	3	1	2	1				
2	26	5.92	55	7	5	9	2	1	5	2	1	0	3



being made with the cooperation of the Department of Supply, the R.A.A.F., and the Meteorological Service. Mr. E. Matthaëi assisted with the reproduction of the mosaic shown in Plate 1.

### Reference

- (1) DILWORTH, C. C., OCCHIALINI, G. P. S., and PAYNE, R. M.—*Nature* **162** : 102-3 (1948).

### Explanation of Plate 1

Mosaic of  $\pi$ -meson coming to rest and disintegrating into a  $\mu$ -meson which in turn comes to rest before disintegrating into an electron at minimum ionizing energy. Additional uncharged particles are emitted but these are not detected by the emulsion. Assuming that conservation of energy and momentum occurs and a single uncharged particle is emitted with the emission of the  $\mu$ -meson the mass of the neutral particle at the  $\pi$ - $\mu$  transformation is less than 10 electron masses and it is generally assumed to be a neutrino.

## A NEW SOURCE OF "L-QUERCITOL" (VIBURNITOL\*)†

By JEAN EWING,‡ G. K. HUGHES,‡ and E. RITCHIE‡

*Stephania hernandifolia* Walp., family Menispermaceae, a slender climber ranging over the coastal districts of eastern Australia, is reputed to be a stock poison(1). It has been examined by Bancroft(2) who isolated a crude alkaloid fraction and by Rennie and Turner(3) who stated that it contained picrotoxin and an alkaloid. In a new investigation of this plant\*\* the presence of picrotoxin could not be demonstrated with certainty but there was readily obtained a crude mixture of alkaloids which has not yet been separated into its constituents. These matters are receiving further attention. However, by concentrating and cooling an alcoholic extract of the leaves a pure crystalline substance was isolated in 0.4 per cent. yield which was identified as "l-quercitol", previously obtained by Power and Tutin(4) from *Gymnema sylvestre* Br.

The substance, which analysed for  $C_6H_{14}O_6$ , contained one molecule of water of crystallization, lost by drying at 118 °C. in a vacuum pistol. It did not react with triphenylmethyl chloride but a penta-acetyl and a pentabenzoyl derivative were obtained by the usual methods. The constants of the substance

\* After this work had been completed and submitted for publication it was announced by Posternak and Schopfer (*Helv. Chim. Acta* **33** : 343 (1950)) that "l-quercitol" and viburnitol are identical and that the substance is 2,3,5/4,6-cyclohexane-pentol. These authors have decided to retain the name viburnitol.

† Manuscript received December 12, 1949.

‡ Department of Organic Chemistry, University of Sydney.

\*\* The plant material was supplied by Mr. L. J. Webb, C.S.I.R.O., and the work is part of the survey of the chemical constituents of the Australian flora being carried out by C.S.I.R.O. and other institutions.

and its two derivatives found in the present work are compared in Table 1 with those reported by Power and Tutin(4). The agreement between some of the figures is excellent but the wide discrepancies in others at first rendered the identification doubtful.

TABLE 1

Present Work	Power and Tutin
<p>"<i>l</i>-Quercitol "</p> <p>m.p. 181.5 °C.</p> <p><math>[\alpha]_D^{20}</math> —50.0° (c, 4%) in water</p> <p>Penta-acetyl "<i>l</i>-quercitol "</p> <p>m.p. 125.5 °C. (a)</p> <p><math>[\alpha]_D^{20}</math> —22.0° (c, 2%) in chloroform</p> <p>Pentabenzoyl "<i>l</i>-quercitol "</p> <p>m.p. 158.5 °C.</p> <p><math>[\alpha]_D^{20}</math> —79.0° (c, 2%) in chloroform</p>	<p>m.p. 174 °C.</p> <p><math>[\alpha]_D</math> —73.9° (c, 4.035%) in water</p> <p>m.p. 124–5 °C.</p> <p><math>[\alpha]_D</math> —26.0 (c, 2.697%) in chloroform</p> <p>m.p. 133 °C. (b) 148 °C. (c)</p> <p><math>[\alpha]_D</math> —79.0° (c, 2.826%) in chloroform</p>

(a) The benzene solvated form, m.p. 87–97 °C., could not be obtained. (b) Amorphous, purified from alcohol. (c) Obtained by adding light petroleum to an alcohol-ethyl acetate solution and then drying at 100 °C. This form could not be prepared.

Fortunately however, a small amount of Power and Tutin's original specimen of "*l*-quercitol " was made available to us for comparison. On re-examination it was found to have m.p. 181.5 °C.,  $[\alpha]_D^{20}$  —50.0° (c, 1%) in water and its penta-acetyl derivative, m.p. 125.5 °C.,  $[\alpha]_D^{20}$  —21.0° (c, 1%) in chloroform. In addition no depressions were observed on taking mixed m.p. of the two substances and of their acetyl derivatives. It must be concluded that the two substances are identical and that the constants given by Power and Tutin are erroneous.

### Experimental

*Isolation of "l-Quercitol "*.—The dried and ground leaves and twigs were exhausted with alcohol at room temperature and the extract concentrated under reduced pressure to a fairly thick syrup. After standing in the refrigerator for several days the "*l*-quercitol " crystallized out. It was collected, washed thoroughly with ether, and recrystallized from alcohol, forming colourless needles, m.p. 181.5 °C. after sintering at about 110 °C. Found: C, 39.2; H, 7.9%. Calculated for  $C_6H_{14}O_6$ : C, 39.5; H, 7.8%. Found: loss of weight on drying at 118 °C. *in vacuo* 9.8%. Calculated for  $C_6H_{12}O_5 \cdot H_2O$ : 9.9%. Found in anhydrous material: C, 43.7; H, 7.5%. Calculated for  $C_6H_{12}O_5$ : C, 43.9; H, 7.4%.

*Penta-acetyl "l-Quercitol "*.—A solution of "*l*-quercitol " (0.5 g.) in pure dry pyridine (8 ml.) containing excess acetic anhydride (6 g.) was refluxed for 1 hour and then poured into water. The product (90% yield) which separated on standing, crystallized from water in colourless needles, m.p. 125.5 °C. (Found: C, 51.3; H, 5.8%. Calculated for  $C_{16}H_{22}O_{10}$ : C, 51.3; H, 5.9%.)

*Pentabenzoyl "l-Quercitol "*.—A solution of "*l*-quercitol " (0.5 g.) and excess benzoyl chloride (3.5 g.) in pure dry pyridine (8 ml.) was refluxed for 1 hour and then diluted with water.

The separated oil gradually solidified. The product (80% yield) was collected, washed thoroughly with water, and recrystallized from alcohol from which it separated as colourless needles, m.p. 158.5 °C. (Found: C, 71.8; H, 4.6%. Calculated for  $C_{41}H_{32}O_{10}$ : C, 71.9; H, 4.7%.)

The authors are grateful to Miss J. Fildes for the analyses; to Dr. S. J. Angyal for advice; and to Dr. A. Albert for the sample of "l-quercitol".

### References

- (1) HURST, E.—"The Poison Plants of New South Wales." p. 120. (The Poison Plants Committee of New South Wales: Sydney, 1942.)
- (2) BANCROFT, J.—*Proc. Linn. Soc. N.S.W.* **14**: 1061 (1889).
- (3) RENNIE, E. H., and TURNER, E. F.—*Trans. Roy. Soc. S. Aust.* **17**: 186 (1893).
- (4) POWER, F. B., and TUTIN, F.—*J. Chem. Soc.* **85**: 624 (1904).

## THE CHEMISTRY OF WESTERN AUSTRALIAN PLANTS\*

### III. URSOLIC ACID FROM *ANTHOTROCHE* SPECIES†

By W. BOTTOMLEY‡ and D. E. WHITE\*\*

Investigation of *Anthotroche* species (family Solanaceae), occurrence of which is confined to South Australia and Western Australia, was undertaken to discover whether they, like the majority of solanaceous plants, contained alkaloids. No alkaloids have been found in the two species examined but one of these, *A. blackii* F. Muell., which occurs in two limited areas near Mullewa in the wheat belt of Western Australia and near the South Australian border, contains a large amount of a triterpene acid and the other, *A. pannosa* Endl., which occurs throughout the low rainfall areas of Western Australia, contains a smaller amount of the same acid. This has been identified as ursolic acid, by comparison of the acid, its acetyl derivative, and its methyl ester with authentic ursolic acid and the corresponding derivatives prepared from authentic material.

Ursolic acid is widely distributed in the plant kingdom and in Australian plants it has been recorded from the Solanaceae in three *Duboisia* species by Trautner and Neufeld(1) and from the Goodeniaceae in *Goodenia ovata* Sm. by Lahey and Strasser(2).

\* Manuscript received May 30, 1950.

† Part II. *Aust. Chem. Inst. J. and Proc.* **17**: 31 (1950).

‡ Present address: Division of Plant Industry, C.S.I.R.O., Canberra, A.C.T.

\*\* Organic Chemistry Department, University of Western Australia, Nedlands, W.A.

### Experimental

All melting points are corrected.

(i) *Extraction*.—Dried, ground leaves and stems of *Anthotroche blackii* collected from the vicinity of Wilroy, W.A., (75 g.) were percolated with cold methanol until the extract was no longer coloured. The extract was then concentrated to a small bulk, water added, and the remainder of the methanol distilled. The precipitated solid was filtered and washed with dilute hydrochloric acid.

The filtrate was carefully made alkaline with sodium hydroxide, extracted with chloroform, and the chloroform solution extracted with dilute hydrochloric acid. This acid extract did not give a precipitate with the usual alkaloid reagents, silicotungstic acid, potassium mercuri-iodide (Mayer's reagent), or iodine in potassium iodide (Wagner's reagent).

The precipitated solid was treated with 5% potassium hydroxide in 50% aqueous alcohol, filtered, and the filtrate acidified with 15% aqueous alcoholic hydrochloric acid. The precipitate (7.1 g.; 9.5%) was repeatedly crystallized from alcohol and formed colourless needles, m.p. 286–287 °C., alone or mixed with authentic ursolic acid purchased from L. Light & Co. Ltd., and purified by repeated crystallization from alcohol.

Found: C, 78.3; H, 10.6%; mol. wt. 439 (Rast) (after drying at 110 °C. *in vacuo*).

Calculated for  $C_{30}H_{48}O_3$ : C, 78.9; H, 10.6%; mol. wt. 456.7.

A sample of *A. pannosa*, collected near Meckering, W.A., (500 g.) and extracted similarly, failed to show alkaloid reactions but yielded ursolic acid (2.5 g.; 0.5%), m.p. 289–291 °C., undepressed by admixture with authentic material.

(ii) *Acetylursolic Acid*.—Ursolic acid (1 g.) was acetylated with acetic anhydride (8 ml.) and anhydrous sodium acetate (2 g.) and the product washed with sodium carbonate, and precipitated from solution in aqueous alcoholic potassium hydroxide with aqueous alcoholic hydrochloric acid. It formed colourless needles from alcohol, m.p. 295–296 °C. alone or mixed with authentic acetylursolic acid.

Found: C, 77.2; H, 10.3%.

Calculated for  $C_{32}H_{50}O_4$ : C, 77.1; H, 10.1%.

(iii) *Methyl Ursolate*.—Ursolic acid (1 g.) was esterified with methyl sulphate (3.0 ml.) and excess aqueous alcoholic potassium hydroxide at 45 °C. The product (0.9 g.) formed long colourless prisms from 75% alcohol. After drying *in vacuo* at 100 °C. it melted at 170–171 °C. alone or on admixture with authentic methyl ursolate.

Found: C, 79.2; H, 10.6%.

Calculated for  $C_{31}H_{50}O_3$ : C, 79.1; H, 10.7%.

Thanks are due to Mr. C. A. Gardner, Government Botanist, for the collection and identification of the material; to the Drug Panel of the Western Australian Department of Industrial Development for a maintenance grant; to the Commonwealth Research Grant to Australian Universities for equipment; and to Dr. Weiler and Dr. Strauss, Oxford, for the microanalyses.

### References

- (1) TRAUTNER, E. M., and NEUFELD, O. E.—*Aust. Chem. Inst. J. and Proc.* **14**: 17 (1947).
- (2) LAHEY, F. N., and STRASSER, P. H. A.—*Aust. Chem. Inst. J. and Proc.* **14**: 432 (1947).





# INTERACTIONS BETWEEN NEUTRON AND PROTON

By P. SWAN\*

[Manuscript received June 22, 1950]

## Summary

An attempt has been made to find an interaction between neutron and proton which will account for not only the binding energy and quadripole moment of the deuteron and the low energy scattering data, but also the results of the experiments on the scattering of 90 MeV. neutrons by protons. Three types of modification of the triplet neutron-proton interaction have been used which embody the following features :

(1) A non-central potential of spherical well form, whose radius of interaction is varied.

(2) A non-central potential whose form is closer to that of the pseudo-scalar meson potential than the usual Rarita-Schwinger form, but which does not possess the objectionable singularities of the former.

(3) The inclusion of a large short-range repulsion.

In each case exchange forces of the usual types have been used. The results obtained, like those of other workers in this field using different forms of interaction, fail to agree with the high energy data.

## I. INTRODUCTION

In view of the unsatisfactory features existing at present in the meson theory of nuclear forces, such as the well-known divergencies and the difficulty of relating the theory to the meson masses and lifetimes found in nature, most recent attempts to derive information about the neutron-proton interaction have proceeded on the basis of an interaction energy of arbitrary form chosen to fit as much of the observed data as possible.

The experimental information falls into two broad categories, the low energy region and the high energy region. In the low energy region, by making use of the known binding energy of the deuteron, we are able to relate the width  $a$  of the potential well to the depth  $V_0$  if this well is represented by a short range interaction potential  $V(r)$  of the form  $V = -V_0$  for  $r < a$  and  $V = 0$  for  $r > a$ , that is, a central spherical well. In point of fact the low energy data have been found to be insensitive to the form of  $V(r)$  so long as it is of short range, and the use of a spherical well facilitates the calculations. As a result of this insensitivity, the experiments on low energy scattering of neutrons by protons do not lead to an accurate value of the range  $a$  for the triplet interaction (parallel spins for the neutron and proton). However, a triplet range of about  $1.6 \times 10^{-13}$  cm. is indicated by the more accurate experiments on the scattering of thermal neutrons by ortho- and para-hydrogen(1), and by crystalline sodium hydride(2), and by the more recent experiments on reflexion from liquid mirrors(3). Experiments at 10.5 MeV. on the scattering of protons by protons

\* Physics Department, University of Melbourne.

appear also to indicate a triplet proton-proton range of  $1.6 \times 10^{-13}$  cm.(4). On the other hand, the singlet range (range between two particles of anti-parallel spins) turns out to have the larger value  $2.8 \times 10^{-13}$  cm., this value being obtained from the accurate low energy proton-proton scattering experiments at 0-3 MeV.

In high energy scattering (10 MeV. or greater), higher values of the orbital angular momentum  $l$  than  $l=0$  will become important, the magnitude and sign of the potentials effective in these interactions of particles for  $l \geq 1$  being governed by the exchange nature of the force. Ordinary (non-exchange) forces do not cause any exchange of coordinates between a neutron and a proton and the interaction is the same for all  $l$ . On the other hand, Majorana exchange forces lead to the exchange of spatial coordinates between two particles, and hence to the reversal of the sign of interaction on passing from even  $l$  to odd  $l$  (5), that is, the  $l=1$  potential will be repulsive. On the basis of meson field theory, ordinary and Majorana forces correspond to neutral meson theory (virtual emission of neutral mesons between particles) and charged meson theory (virtual emission of charged mesons between particles) respectively. The neutral theory is inadmissible, as it will not lead to the observed saturation of nuclear forces(5), and furthermore gives quite the wrong form of angular distribution for the scattering of neutrons by protons at 90 MeV.(6) and the scattering of protons by protons at 10 and 14.5 MeV.(7, 8). The charged theory, although yielding saturation of nuclear forces, suffers from the objection that it will (at least in the first approximation) give zero forces between two protons, that is, the charged interaction is not charge-independent (i.e. the same for  $p-p$  as for  $n-p$ ) as indicated by experiment. If neutral and charged theory are mixed in a certain proportion, the resultant potential (symmetric theory) is charge independent and leads to saturation of nuclear forces. However, while all the above potentials are consistent with the results of scattering experiments up to about 20 MeV. none of them give total cross sections or angular distribution curves in agreement with experiments at higher energies(6), the calculated values of the total cross section  $\sigma$  and the anisotropy ratio ( $d\sigma_{180}/d\sigma_{90}$ ) being too large. Nevertheless the charged and symmetric theories are in much better qualitative agreement with experiment than the neutral theory, as they both predict the typical peak of backward scattering (c.m. system).

Experimentally it has been found that the deuteron possesses a small electric quadrupole moment, corresponding to a non-spherical charge distribution prolate along the spin axis. The  $n-p$  force therefore does not lie entirely along the line joining them, that is, the interaction is non-central like that between two dipoles. The inclusion of non-central forces has been found to give results for the scattering of neutrons by protons almost identical with those for pure central forces at low energies, but at high energies, say 90 MeV., the calculated scattering disagrees with the experimental value of the total cross section  $\sigma=0.082 \times 10^{-24}$  cm.<sup>2</sup> and the anisotropy ratio  $\chi(d\sigma_{180}/d\sigma_{90})=3$ . In fact the inclusion of non-central forces worsens the agreement at high energies, as it usually results in values of  $\sigma$  at least 50 per cent. too large and of  $\chi$  from 6 upwards. In the present work an attempt is made to obtain agreement with

experiment in the high energy region (*a*) by taking different ranges for the singlet and triplet interactions instead of identical ranges as had previously been assumed, and (*b*) by taking radically different forms of non-central and central interaction to those previously used.

## II. THE RARITA-SCHWINGER NON-CENTRAL INTERACTION

The theory of non-central or tensor forces was developed by Rarita and Schwinger(9), who avoided the inadmissible  $1/r^3$  singularity occurring in all non-central meson potentials by postulating a linear combination of central and non-central forces for the interaction operator, viz.

$$V = -[1 - \frac{1}{2}g + \frac{1}{2}g\sigma_1 \cdot \sigma_2 + \gamma S_{12}]J(r),$$

where  $\sigma_1$ ,  $\sigma_2$  are the nucleonic spin operators for neutron and proton,

$g$ ,  $\gamma$  are arbitrary constants,

$$S_{12} = \left( 3 \frac{\sigma_1 \cdot r \sigma_2 \cdot r}{r^2} - \sigma_1 \cdot \sigma_2 \right).$$

$S_{12}J(r)$  is the non-central part of the potential and is due to the coupling of the  $^3S_1$  ground state with a small amount of  $^3D_1$  state (non-central). For the singlet state this potential reduces to the purely central form  $-(1-2g)J(r)$ , while for the triplet state it becomes  $-(1+\gamma S_{12})J(r)$ . Rarita and Schwinger were able to solve the ground state problem of the deuteron, that is, calculate values of the potential well parameters which gave the observed values of the deuteron binding energy quadrupole moment, and percentage of time (4 per cent.) spent by the deuteron in the  $D$  state. The results obtained by them for the scattering of neutrons by protons at energies up to 15 MeV.(10), and later by Hu and Massey(11) up to 20 MeV. were in good agreement with experiment, but at higher energies serious disagreement appeared, as mentioned earlier.

## III. GROUND STATE PROBLEM OF THE DEUTERON WITH UNEQUAL SINGLET AND TRIPLET RANGES

We propose to try to obtain the correct binding energy and quadrupole moment of the deuteron using a singlet spherical well of range  $2.8 \times 10^{-13}$  cm. (as required by low energy proton-proton scattering) and a triplet spherical well of range  $1.6 \times 10^{-13}$  cm. (as obtained from the slow neutron scattering experiments). Two cases were considered:

- (1) Central and non-central triplet wells with a common range  $1.6 \times 10^{-13}$  cm.
- (2) A central triplet well of range  $1.6 \times 10^{-13}$  cm. with a non-central triplet well with a larger range, say  $2.8 \times 10^{-13}$  cm.

Case (2) was tried when the first case failed to give even a positive quadrupole moment.

Case (1) involves repetition of Rarita and Schwinger's calculation(9) for a smaller well range. A triplet state with total angular momentum  $J=1$  is, by the rules for compounding angular momenta, a mixture of  $^3S_1$ ,  $^3P_1$ , and  $^3D_1$  states and introduction of the quantum number parity  $(-1)^l$  enables further classification into an even state  $^3S_1 + ^3D_1$  and an odd state  $^3P_1$ . As for central



forces it is known that the  ${}^3S_1$  state is the ground state then it follows that for non-central forces the ground state will be  ${}^3S_1 + {}^3D_1$ , where the deuteron spends only a small percentage of time in the  ${}^3D_1$  state. One obtains for the wave function of the  ${}^3S_1 + {}^3D_1$  states (9)

$$\psi = (4\pi)^{-\frac{1}{2}} \left( \frac{u}{r} + 2^{-3/2} S_{12} \frac{w}{r} \right) \chi_1^m,$$

where  $u/r$ ,  $w/r$  are the  ${}^3S_1$  and  ${}^3D_1$  radial wave functions respectively.

$$S_{12} = \left( 3 \cdot \frac{\sigma_1 \cdot r \sigma_2 \cdot r}{r^2} - 1 \right),$$

$\chi_1^m$  = spin wave function for magnetic quantum number " $m$ ".

Substituting  $\psi$  and the Rarita-Schwinger non-central potential

$$V = -(1 - \frac{1}{2}g + \frac{1}{2}g\sigma_1 \cdot \sigma_2 + \gamma S_{12})J(r)$$

into the Schrodinger wave equation leads to two second order, coupled differential equations in  $u$  and  $w$ .

$$\begin{aligned} \frac{d^2 u}{dr^2} + \frac{4\pi^2 M}{\hbar^2} (E + J) u &= -2 \frac{4\pi^2 M}{\hbar^2} J w, \\ \frac{d^2 w}{dr^2} - \frac{6w}{r^2} + \frac{4\pi^2 M}{\hbar^2} [E + (1 - 2\gamma)J] w &= -2^{3/2} \frac{4\pi^2 M}{\hbar^2} J u. \end{aligned}$$

These equations may be solved by expanding both  $u$  and  $w$  in an infinite power series, then using the boundary conditions for continuity of  $u$  and  $w$  across the nuclear boundary, and finally endeavouring simultaneously to satisfy the normalization condition and obtain the correct values of the binding energy and the quadrupole moment. The method of treatment is that given below for the more general case (2). The detailed calculations showed that if the correct binding energy of 2.185 MeV. is obtained, it is not possible to obtain a positive quadrupole moment.

Case (2), where unequal ranges of central and non-central triplet wells are assumed, must therefore be considered. It is more difficult in detail than case (1). We write the triplet potential  $J$  as follows:

$$J = -\{J_1(r) + \gamma S_{12} J_2(r)\},$$

where the central and non-central potentials  $J_1$  and  $J_2$  are taken as spherical wells of the following dimensions:

$$\begin{aligned} J_1(r) &= -V_1 \text{ for } 0 < r < r_1, \\ &= 0 \text{ for } r_1 < r; \\ J_2(r) &= -V_2 \text{ for } 0 < r < r_2, \\ &= 0 \text{ for } r_2 < r, \end{aligned}$$

where  $r_1 = 1.6 \times 10^{-13}$  cm.,  $r_2 = 2.8 \times 10^{-13}$  cm.

Taking the ground state as a strongly coupled mixture of  ${}^3S_1$  and  ${}^3D_1$  states, the wave equations are found to be

$$\begin{aligned} \frac{d^2 u}{dr^2} + \frac{4\pi^2 M}{\hbar^2} [E + J_1(r)] u &= 2^{3/2} \frac{4\pi^2 M}{\hbar^2} J_2(r) w, \\ \frac{d^2 w}{dr^2} - \frac{6w}{r^2} + \frac{4\pi^2 M}{\hbar^2} [E + J_1(r) - 2\gamma J_2(r)] w &= -2^{3/2} \frac{4\pi^2 M}{\hbar^2} J_2(r) u, \end{aligned}$$

or more concisely,

$$\left(\frac{d^2}{dr^2} + \kappa^2\right)u = -\lambda^2 w,$$

$$\left(\frac{d^2}{dr^2} - \frac{6}{r^2} + \kappa'^2\right)w = -\lambda^2 u,$$

where

$$\kappa^2 = \frac{4\pi^2 M}{h^2} (V_0 - |E_0|),$$

$$\kappa'^2 = \frac{4\pi^2 M}{h^2} [(V_1 - 2V_2) - |E_0|],$$

$$\lambda^2 = 2^{3/2} \gamma \frac{4\pi^2 M}{h^2} V_2.$$

The solutions of these coupled differential equations must be consistent with the following conditions :

(a)  $E_0 = 2.185$  MeV.,

(b) Electric quadripole moment  $Q = \frac{2}{10} \int_0^\infty r^2 (uw - 2^{-3/2} w^2) dr,$

(c) Normalization condition  $\int_0^\infty (u^2 + w^2) dr = 1,$

from which we can determine the values (if any) of the three unknowns  $V_1$ ,  $V_2$ , and  $\gamma$  which satisfy these conditions.

There are four possible series solutions to the differential equations for  $u$ ,

$$u_1 = \sum_{n=0}^\infty a_n r^{n+5}, \text{ regular at the origin,}$$

$$u_2 = \sum_{n=0}^\infty c_n r^n, \text{ regular at the origin,}$$

$$u_3 = \ln r \sum_{n=0}^\infty a_n r^{n+5} + \sum_{n=0}^\infty b_n r^{n+1}, \text{ regular at the origin,}$$

$$u_4 = \ln r \sum_{n=0}^\infty c_n r^n + \sum_{n=0}^\infty d_n r^{n+4}, \text{ irregular solution.}$$

The most general possible solution is a linear combination of these, and  $w$  can be obtained by direct substitution into the second differential equation.

We divide the region into three ranges :

$$(1) \quad 0 < r < r_1,$$

$$(2) \quad r_1 < r < r_2,$$

$$(3) \quad r > r_2.$$

In range (1) the general solution is not required as the condition that  $u$ ,  $w$  be zero at  $r=0$  rules out  $u_4$ , and provides a relation to eliminate one of the remaining series. One then finds

$$u(r < r_1) = \sum_{n=0}^\infty a_n r^{n+1} + \ln r \sum_{n=0}^\infty c_n r^{n+2},$$

$$w(r < r_1) = \sum_{n=0}^\infty b_n r^{n+3} + \ln r \sum_{n=0}^\infty d_n r^{n+3}.$$

Recurrence relations between the series constants can then be found by substitution in the differential equations, reducing terms to dependence on  $a_0$  and  $b_0$ .

In range (2), where  $V_1=0$  but  $V_2 \neq 0$ , the general solution is required. We have\*

$$u(r_1 < r < r_2) = u_1 + u_2 + u_3 + u_4,$$

$$w(r_1 < r < r_2) = w_1 + w_2 + w_3 + w_4,$$

where

$$w_1 = -[1/(\lambda r)^2] \cdot \left[ \sum_{n=0}^{\infty} (n+4)(n+5)a_n r^{n+3} + (\lambda r)^2 \sum_{n=0}^{\infty} a_n r^{n+5} \right],$$

$$w_2 = -[1/(\lambda r)^2] \left[ \sum_{n=0}^{\infty} n(n-1)c_n r^{n-2} + (\lambda r)^2 \sum_{n=0}^{\infty} c_n r^n \right],$$

$$w_3 = -[1/(\lambda r)^2] \cdot \left[ \ln r \sum_{n=0}^{\infty} (n+4)(n+5)a_n r^{n+3} + \sum_{n=0}^{\infty} (2n+9)a_n r^{n+3} \right. \\ \left. + (\lambda r)^2 \sum_{n=0}^{\infty} b_n r^{n+1} + \sum_{n=0}^{\infty} n(n+1)b_n r^{n-1} \right],$$

$$w_4 = -[1/(\lambda r)^2] \cdot \left[ \ln r \sum_{n=0}^{\infty} n(n-1)c_n r^{n-2} + \sum_{n=0}^{\infty} (2n-1)c_n r^{n-2} \right. \\ \left. + (\lambda r)^2 \sum_{n=0}^{\infty} d_n r^{n+4} + \sum_{n=0}^{\infty} (n+3)(n+4)d_n r^{n+2} \right].$$

In range (3),  $V_1 = V_2 = 0$  ;

then

$$u(r > r_2) = A e^{-\alpha r},$$

$$w(r > r_2) = B e^{-\alpha r} \left[ 1 + \frac{3}{\alpha r} + \frac{3}{(\alpha r)^2} \right].$$

The eight constants involved in the series and the unknown constants  $V_1$  and  $V_2$  may be determined in terms of  $\gamma$  by using the boundary conditions that  $u$ ,  $w$  and their first derivatives be continuous at  $r=r_1$  and  $r=r_2$  and satisfy the quadrupole moment and normalization integrals (eight conditions). To simplify the numerical work, a number of pairs of trial solutions for  $V_1$  and  $\gamma V_2$  were chosen and in each case the resulting eight linear simultaneous equations were solved partly by direct elimination of constants and partly by the use of determinants. The quadrupole moment then turns out to be negative for all choice of  $V_1$  and  $\gamma V_2$ , in contradiction to experimental evidence.

However, Padfield(12) explored a similar field with a differential analyser and showed that if, while retaining the central triplet range  $r_1 = 1.6 \times 10^{-13}$  cm., the non-central range was increased still further to  $r_2 = 3.07 \times 10^{-13}$  cm., the correct positive quadrupole moment could be obtained. Castillejo and Richardson(13) used these constants to calculate the scattering of neutrons by protons at 83 MeV., but found that the scattering cross section and anisotropy ratio obtained were still rather larger than the experimental values.

Actually at high energies, say 90 MeV., the inclusion of non-central forces tends to worsen the agreement with experiment ; thus calculations by the writer

\* These expressions for  $u$  and  $w$  are linear combinations of the four component solutions, since the initial coefficients  $a_0$ ,  $b_0$ ,  $c_0$ , and  $d_0$  are arbitrary.

at 90 MeV. using the symmetrical and charged exchange theories with purely central forces (i.e. neglecting the small electric quadripole moment) and assuming triplet and singlet ranges of  $1.6 \times 10^{-13}$  cm. and  $2.80 \times 10^{-13}$  cm. respectively yields

(a) symmetric theory :  $\sigma = 0.104 \times 10^{-24}$  cm.<sup>2</sup>, anisotropy ratio  $\chi = 8.1$  ;

(b) charged theory :  $\sigma = 0.083 \times 10^{-24}$  cm.<sup>2</sup>,  $\chi = 5.2$ .

Both sets of values, especially those for the charged theory, are in much better agreement with experiment than the corresponding non-central calculations.  $\chi$  is, however, too large.

#### IV. GROUND STATE PROBLEM AND SCATTERING WITH AN EMPIRICALLY MODIFIED PSEUDO-SCALAR POTENTIAL

Owing to the inadmissible singularity at the origin ( $\sim 1/r^3$ ) involved in any meson theory potential, the non-central potentials that have so far been examined can be divided into two classes :

(a) The Rarita-Schwinger empirical potential which is taken arbitrarily as a linear combination of central and non-central terms :

$$V = -[1 - \frac{1}{2}g + \frac{1}{2}g\sigma_1 \cdot \sigma_2 + \gamma S_{12}]J(r).$$

(b) The meson field theory potentials with cut-off, e.g. the pseudo-scalar theory leads to

$$V = +g^2 \left[ \sigma_1 \cdot \sigma_2 + \left( 1 + \frac{3}{\mu r} + \frac{3}{(\mu r)^2} \right) S_{12} \right] \mu^2 \frac{e^{-\mu r}}{r},$$

where  $\mu = mc/(\hbar/2\pi)$ , and  $m$  = meson mass. The  $1/r^3$  term makes it impossible to obtain a solution of the wave equation regular at the origin, so that one obtains an infinite binding energy for the deuteron, but this difficulty may be removed by postulating that the potential has some finite value such as zero inside a given radius. However, calculation has shown that this cut-off radius would have to be much the same as the range of the potential, and hence the cut-off method must be rejected.

It was thought of interest to try an empirical potential of closely analogous form to the pseudo-scalar potential, but without an inadmissible singularity, viz.

$$V = - \left[ 1 - \frac{1}{2}g + \frac{1}{2}g\sigma_1 \cdot \sigma_2 + \gamma \left( 1 + \frac{3}{\mu r} + \frac{3}{(\mu r)^2} \right) S_{12} \right] J(r),$$

where

$$\begin{aligned} J(r) &= V_0 \text{ for } r < 2.8 \times 10^{-13} \text{ cm.}, \\ &= 0 \text{ for } r > 2.8 \times 10^{-13} \text{ cm.} \end{aligned}$$

Since no terms occur in  $V$  of higher singularity than  $1/r^2$ , the differential equation will have a solution regular at the origin ; but owing to the greater complexity of this potential, it is no longer convenient to use the series method for solution of the differential equations involved. If  $u/r$  and  $w/r$  denote the  $^3S_1$  and  $^3D_1$  radial wave functions as before, we have

$$\begin{aligned} \left( \frac{d^2}{dr^2} + \kappa^2 \right) u &= -2^{3/2} \left( 1 + \frac{3}{\mu r} + \frac{3}{(\mu r)^2} \right) \lambda^2 w, \\ \left[ \frac{d^2}{dr^2} - \frac{6}{r^2} - 2\lambda^2 \left( 1 + \frac{3}{\mu r} + \frac{3}{(\mu r)^2} \right) + \kappa^2 \right] w &= -2^{3/2} \left( 1 + \frac{3}{\mu r} + \frac{3}{(\mu r)^2} \right) \lambda^2 u. \end{aligned}$$



These coupled equations in  $u$  and  $w$  may be solved numerically in a step-by-step process, in which each new value of  $u$  for a given finite step in  $r$  obtained from the first equation is used in the second equation to give a new value of  $w$  for the same value of  $r$ . The value of  $u$  is then extended over another step in  $r$ , and so on.

The method employed for the solution of the ground state problem of the deuteron involved using numerical integration as a semi-algebraic process in which two unknown constants were retained throughout the progressive outward steps and eliminated by fitting of the continuity conditions at the edge of the well. The finite difference formulae used (14, see equation 4.13) for equations of the type

$$\begin{aligned}u'' + g(u, r) &= 0, \\w'' + h(w, r) &= 0,\end{aligned}$$

were

$$\begin{aligned}u_1 &= 2u_0 - h^2 g_0 - u_{-1} + \Delta, \\w_1 &= 2w_0 - h^2 h_0 - w_{-1} + \Delta',\end{aligned}$$

where

$$\Delta = \frac{1}{24} h^4 u_0 \quad \text{and} \quad \Delta' = \frac{1}{24} h^4 w_0,$$

Only four steps were needed as the  $\Delta$  and  $\Delta'$  were sufficiently small to keep the total error involved less than about two per cent. As two of the boundary conditions involve the first derivatives  $u'$  and  $w'$ , these latter were found by Taylor's theorem: thus for

$$h = \frac{1}{4}, \quad u_3 = u_4 - \frac{1}{4}u'_4 + \frac{1}{32}u''_4 - \frac{1}{384}u'''_4 + \frac{1}{6,144}u^{IV}_4 \dots,$$

where  $u_3, u_4$  are known and  $u''_4, u'''_4$ , etc. are found in terms of  $u_4$  and  $u'_4$  by the original differential equations.

Finally one obtains from the numerical integration the values of  $u_4$  and  $w_4$  in terms of  $u_1$  and  $w_1$ :

$$\begin{aligned}u_4 &= A_4 u_1 + B_4 w_1, \\w_4 &= C_4 u_1 + D_4 w_1,\end{aligned}$$

and combining these relations with the continuity conditions for  $u, w, u'$ , and  $w'$  at the nuclear boundary, one obtains the following two equations:

$$\begin{aligned}u_1(m_1 A_4 + n_1 C_4 - A_3) - w_1(m_1 B_4 + n_1 D_4 - B_3) &= 0, \\u_1(m_1 A_4 + n_1 C_4 - C_3) - w_1(m_1 B_4 + n_1 D_4 - D_3) &= 0,\end{aligned}$$

and if these are to be mutually consistent, one must have

$$(M_1 A_4 + N_1 C_4 - A_3)(m_1 B_4 + n_1 D_4 - D_3) = (m_1 A_4 + n_1 C_4 - C_3)(M_1 B_4 + N_1 D_4 - B_3),$$

where  $A_4, B_4, C_4, D_4, M_1$ , etc. are complicated polynomials in  $V_0$  and  $\gamma V_0$ . This relation between  $V_0$  and  $\gamma V_0$  may be solved by trial and error for  $V_0$  in terms of  $\gamma V_0$  and the resultant trial solutions substituted in the quadripole moment integral until the correct combination is found.

The solution for range  $2.8 \times 10^{-13}$  cm. yielding the correct binding energy and quadripole moment is  $V_1 = 0$ ,  $V_2 = 3.31$  MeV., that is, a pure non-central force. In the calculation a meson mass  $m = 300 m_e$  ( $m_e$  = electron mass) was used in compliance with cosmic ray evidence.

This non-central potential was used to calculate the scattering of neutrons by protons at 90 MeV. by the accurate phase-shift method(10). The total cross section turned out large as usual, viz.  $\sigma=0.20$  barns, the *P*-wave scattering being especially large. This large cross section appears to be typical of a purely non-central potential and rules out all chance of agreement with experiment.

#### V. NUCLEAR POTENTIAL WITH A SHORT RANGE, ALMOST INFINITE, REPULSION

Morse, Schiff, and Fisk(15) suggested that, analogous to the chemical binding force between molecules, the nuclear potential might be strongly repulsive at distances small compared with the potential range. They took  $V = -2De^{2/r_0(r_1-r)} + De^{4/r_0(r_1-r)}$ , showing that this would satisfactorily explain the saturation of nuclear forces and low energy neutron-proton scattering up to

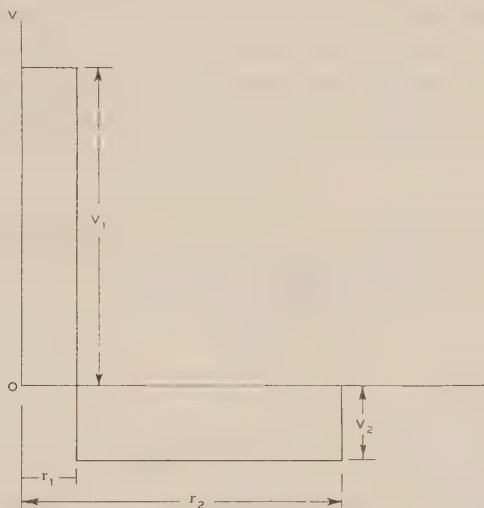


Fig. 1

several MeV. Since the treatment of this potential must be mainly numerical (in general no analytic solution of the wave equation exists), it is simpler to consider a spherical well potential with a short range spherical repulsion (Fig. 1). An additional simplification is to take the barrier as almost infinite in height, as any results for one of finite magnitude will be intermediate between the two limiting cases of the infinite barrier and the simple potential well.

Take

$$\left. \begin{aligned} V(r < r_1) &= V_1, \\ V(r_1 < r < r_2) &= -V_0, \\ V(r > r_2) &= 0, \end{aligned} \right\} \text{ with } r_0 = r_2 - r_1.$$

By utilizing the continuity conditions for the radial wave function and its first derivative at  $r=r_1$  and  $r=r_2$ , one may easily show that

$$\tan k_0 r_0 = -\frac{k_0 r_0}{\alpha r_0} + \frac{1}{\alpha r_0} \left( \frac{\tanh k_1 r_1}{k_1 r_0} \right) (k_0 r_0) [(k_0 r_0) \tan (k_0 r_0) - (\alpha r_0)],$$

where

$$\left. \begin{aligned} k_1 &= \sqrt{\frac{4\pi^2 M}{h^2} (W + V_1)}, \\ k_0 &= \sqrt{\frac{4\pi^2 M}{h^2} (V_0 - W)}, \\ \alpha &= \sqrt{\frac{4\pi^2 M}{h^2} W}, \end{aligned} \right\} \quad \text{and } E = -W \text{ is the binding energy of the deuteron.}$$

In the limit  $V_1 \rightarrow \infty$ , this yields

$$\tan k_0 r_0 = -\frac{k_0 r_0}{\alpha r_0},$$

which is the same relation between magnitude and range as for a simple potential well of range  $r_0$  and depth  $V_0$ . Thus if we take  $r_0 = 2.8 \times 10^{-13}$  cm. as usual, the infinite barrier does not alter the binding energy of the deuteron.

To determine the range  $r_1$  of the infinite barrier use may be made of the experiments on the scattering of thermal neutrons by ortho- and para-hydrogen. These yield a triplet scattering amplitude  $a_1 = 0.516 \times 10^{-12}$  cm. at zero incident neutron energy ( $a_1 = (\sin \delta)/k$ , where  $\delta$  = triplet phase shift,  $k = \sqrt{4\pi^2 ME/h^2}$ ,  $E$  = incident neutron energy). Solution of the scattering problem for  $l=0$  yields

$$\delta = \delta_{r_0} - k r_1,$$

where  $\delta_{r_0}$  is the phase shift for the well alone and  $\delta$  is the total phase shift for well and barrier; hence at zero incident energy,

$$a_1 \rightarrow \frac{\alpha}{k} = \frac{\delta_{r_0}}{k} - r_1.$$

Use of the relation (16)

$$\alpha^2 a_1^2 = 1 + (\alpha r_0) + 0.3447(\alpha r_0)^2 + 0.0246(\alpha r_0)^3 - 0.0117(\alpha r_0)^4 + \dots,$$

for a simple spherical potential well range  $2.8 \times 10^{-13}$  cm. yields  $a_1 = 5.85 \times 10^{-13}$  cm. Hence a spherical triplet well range  $r_0 = 2.8 \times 10^{-13}$  cm., plus an infinite barrier range  $r_1 = 0.69 \times 10^{-13}$  cm. will give the experimental value  $a_1 = 5.16 \times 10^{-13}$  cm., this being an alternative to the case of no barrier where it is necessary to take the triplet range  $1.6 \times 10^{-13}$  cm. to obtain  $a_1 = 5.16 \times 10^{-13}$  cm.

To evaluate the scattering of neutrons by protons for ordinary forces (neutral meson theory), we see that as an infinite barrier is impenetrable to the incident wave, the latter suffers a phase shift  $\sigma_l$  given by  $J_{l+1/2}(k r_1 + \sigma_1) = 0$  in addition to the phase shift due to the well,  $J_{l/2}(k r)$  being the plane-wave solution of order  $l$  to the wave equation.

For

$$l=0; \quad \sin(k r_1 + \sigma_0) = 0 \quad \text{and hence } \sigma_0 = -k r_1;$$

$$l=1; \quad \frac{\sin(k r_1 + \sigma_1)}{k r_1} - \cos(k r_1 + \sigma_1) = 0;$$

therefore  $\sigma_1 = \text{artan}(k r_1) - k r_1 \sim -\frac{1}{3}(k r_1)^3$  for  $(k r_1)$  small;

similarly  $l=2$  yields  $\sigma_2 = \text{artan} \left[ \frac{3(k r_1)}{3 - (k r_1)^2} \right] - k r_1 \sim -\frac{2}{15}(k r_1)^5$  for  $(k r_1)$  small.

Note that whereas  $\sigma_0$  is a comparatively large correction to be subtracted from the spherical well phase for  $l=0$ ,  $\sigma_1$ ,  $\sigma_2$ , etc. are almost zero for energies up to about 30 MeV.

First consider the zero energy scattering of neutrons by protons. For a simple potential well one has the total cross section for scattering(5)

$$=4\pi\left[\frac{3}{4\alpha^2}(1+\frac{1}{2}\alpha r_0)^2+\frac{1}{4\alpha_s^2}(1-\frac{1}{2}\alpha_s r_0)^2\right],$$

where

$$\alpha = \sqrt{\frac{4\pi^2 M}{\hbar^2} W} \quad \text{and} \quad -W = \text{binding energy of deuteron (triplet state)} \\ = 2.185 \text{ MeV.}$$

$$\alpha_s = \sqrt{\frac{4\pi^2 M}{\hbar^2} W_s} \quad \text{and} \quad W_s = \text{energy of virtual level (singlet state)} \\ = 0.064 \text{ MeV.}$$

Introduction of an infinite barrier range  $r_1$  modifies this to

$$\sigma = 4\pi \left[ \frac{3}{4\alpha^2} (1 + \frac{1}{2}\alpha r_0 - \alpha r_1)^2 + \frac{1}{4\alpha_s^2} (1 - \frac{1}{2}\alpha_s r_0 - \alpha_s r_1)^2 \right].$$

Numerical substitution reduces the theoretical  $\sigma$  from about  $20.5 \times 10^{-24} \text{ cm.}^2$  to  $19.9 \times 10^{-24} \text{ cm.}^2$  as compared with the experimental value  $20 \times 10^{-24} \text{ cm.}^2$ ; thus the barrier makes no significant difference to the zero energy scattering. At medium energies, e.g. 10 MeV., triplet phase  $\delta_T$

$$\delta_T = \delta_{r_0} - kr_1 = 94^\circ 58' - 13^\circ 44' = 81^\circ 14' \quad \text{and singlet phase } \delta_s \\ \delta_s = \delta_{r_0} - kr_1 = 84^\circ 46' - 13^\circ 44' = 71^\circ 2'.$$

Then

$$\sigma = \frac{4\pi}{k^2} \left( \frac{3}{4} \sin^2 \delta_T + \frac{1}{4} \sin^2 \delta_s \right) = 0.4264 \times 10^{-24} \text{ cm.}^2,$$

which compares with the value

$$\sigma = 0.4425 \times 10^{-24} \text{ cm.}^2$$

using the spherical well potential only; thus the barrier makes only a negligible difference at this energy. However, for higher energies, say 90 MeV.,  $\delta_T$  and  $\delta_s$  are decreased to a negligible value while the magnitude of the odd (negative) phases  $\delta_1$  is increased considerably and this counterbalancing effect by the higher phases prevents the total cross section for scattering from being appreciably reduced as required to give agreement with experiment. Thus ordinary forces with a barrier cannot solve the high energy scattering difficulties.

Exchange forces change the high energy situation entirely as for odd  $l$  the barrier is converted into a deep well and the spherical well into a shallow barrier. As the existence of resonance levels for scattering must now be expected, it is necessary to limit the inverted barrier height to any one of a discrete set of large values.

The only experiments offering clear-cut information on the  $l=1$  component are those of Wilson(7) on the scattering of protons by protons at 10 MeV. The



phase value  $K_1 = -0.4^\circ$  obtained(4) may be said to determine the barrier height. Fitting boundary conditions as usual, at  $r=r_1$  and  $r=r_2$

at  $r_1$

$$\frac{k_1 r_1 \sin(k_1 r_1)}{\sin(k_1 r_1)/k_1 r_1 - \cos(kr_1)} = 1 + \left\{ \frac{e^{\kappa_2 r_1}(\kappa_2 r_1 - 1 - 1/\kappa_2 r_1) - \gamma e^{-\kappa_2 r_1}(\kappa_2 r_1 + 1 + 1/\kappa_2 r_1)}{e^{\kappa_2 r_1}(1 - 1/\kappa_2 r_1) + \gamma e^{-\kappa_2 r_1}(1 + 1/\kappa_2 r_1)} \right\}$$

at  $r_2$

$$\begin{aligned} \frac{F'_i}{F_i} &= \frac{\Phi_1^*}{\Phi_1} - 3/(\Phi_1 \Theta_1 + 3C_1^2 \rho^3 \Phi_1^2 \cot k_1) \\ &= \left\{ \frac{e^{\kappa_2 r_2}(\kappa_2 r_2 - 1 - 1/\kappa_2 r_2) - \gamma e^{-\kappa_2 r_2}(\kappa_2 r_2 + 1 + 1/\kappa_2 r_2)}{e^{\kappa_2 r_2}(1 - 1/\kappa_2 r_2) + \gamma e^{-\kappa_2 r_2}(1 + 1/\kappa_2 r_2)} \right\} \end{aligned}$$

using the notation of Yost, Wheeler, and Breit(17); also  $\gamma$  is an arbitrary constant,

$$k_1 = \sqrt{4\pi^2 M(E + V_1)/\hbar^2}.$$

Substituting  $K_1 = -0.4^\circ$ , one finds

$$\frac{(k_1 r_1)^2 \tan k_1 r_1}{\tan k_1 r_1 - k_1 r_1} = +0.28727,$$

and hence  $k_1$ , or ultimately the barrier height  $V_1$ , must be chosen to comply with this condition. The solution is multi-valued and a typical large value would be  $x = 100\pi - 0.0009144$  at 10 MeV. (lab.) for  $p$ - $p$  scattering. As long as a large  $V_1$  is taken, the number of multiples of  $\pi$  included in  $x$  will be immaterial.

Applying this to neutron-proton scattering at 90 MeV. (exchange forces) one finds, after allowing for the subtraction of the Coulomb field and also for the different incident energy, that  $k_1 r_1 = 100\pi + 0.05221$ . Calculation of the scattering shows that  $\delta_0$  is reduced considerably, but  $\delta_1$  (negative) is increased in the negative sense and the resultant effect is to reduce the total cross section only slightly and also to increase the anisotropy ratio ( $d\sigma_{180}/d\sigma_{90}$ ). Thus agreement with experiment is not facilitated and a repulsive barrier cannot explain high energy scattering even when exchange forces are included. However, it is compatible with all low energy data.

This failure, coupled with that of other workers in this field using still other forms of interaction, leads one to suspect that it is not possible to represent the interaction between nuclear particles at high energies by a simple static potential.

## VI. ACKNOWLEDGMENT

The writer wishes to thank Associate Professor C. B. O. Mohr for his advice throughout this work and in particular for suggesting investigation of the repulsive barrier.

## VII. REFERENCES

- (1) SUTTON, R. B., MCDANIEL, B. D., ANDERSON, E. E., and LAVATELLI, L. S.—*Phys. Rev.* **72**: 1147 (1947).
- (2) SHULL, C. G., WOLLAN, E. O., MORTON, G. A., and DAVIDSON, W. L.—*Phys. Rev.* **73**: 842 (1948).

- (3) HUGHES, D. J., BURG, M. T., and RINGO, G. R.—*Phys. Rev.* **77** : 291 (1950).
- (4) SWAN, P.—*Phys. Rev.* **77** : 288 (1950).
- (5) BETHE, H. A.—“Elementary Nuclear Theory.” (John Wiley & Sons : New York, 1942.)
- (6) COOK, L. J., McMILLAN, E. M., PETERSON, J. M., and SEWELL, D. C.—*Phys. Rev.* **72** : 1264 (1947).
- (7) WILSON, R. R.—*Phys. Rev.* **71** : 384 (1947).
- (8) WILSON, R. R., LOFGREN, E. J., RICHARDSON, J. R., WRIGHT, B. T., and SCHANKLAND, R. S.—*Phys. Rev.* **72** : 1264 (1947).
- (9) RARITA, W., and SCHWINGER, J.—*Phys. Rev.* **59** : 436 (1941).
- (10) RARITA, W., and SCHWINGER, J.—*Phys. Rev.* **59** : 556 (1941).
- (11) HU, TSI-MING, and MASSEY, H. S. W.—*Proc. Roy. Soc. A* **196** : 135 (1949).
- (12) PADFIELD, D.—*Nature* **163** : 22 (1949).
- (13) CASTILLEJO, L., and RICHARDSON, H. T.—*Phys. Rev.* **76** : 1732 (1949).
- (14) FOX, L., and GOODWIN, E. T.—*Proc. Camb. Phil. Soc.* **45** : 373 (1949).
- (15) MORSE, P. M., SCHIFF, J. B., and FISK, L. I.—*Phys. Rev.* **50** : 748 (1936).
- (16) ROSENFELD, L.—“Nuclear Forces.” p. 76. (North Holland Publishing Co. : Amsterdam, 1948.)
- (17) YOST, F. L., WHEELER, J. A., and BREIT, G.—*Phys. Rev.* **49** : 174 (1936).

# GEOMETRICAL OPTICS OF CONCAVE MIRRORS AND OF COMBINATIONS OF MIRRORS

By J. SHEARER\*

[Manuscript received July 20, 1950]

## Summary

A study of the geometrical optics of a concave spherical mirror is made over the range of incidence from normal to grazing.

An examination is made of methods of correcting for astigmatism at oblique incidence by the use of "crossed" mirrors (spherical and cylindrical), of "parallel" cylindrical mirrors, and of a single mirror with different radii of curvature in planes at right angles. Such a single mirror combines in one piece image formation and beam deviation.

## I. INTRODUCTION

The optical properties of concave spherical mirrors for small objects located on or near the axis are discussed in any elementary text. Until Kirkpatrick and Baez† developed an X-ray microscope employing concave spherical mirrors no study seems to have been made of the optical properties of such mirrors (and of combinations of such mirrors) for objects remote from the axis. It is of interest to examine more generally than do Kirkpatrick and Baez in the publication cited the geometrical optics of concave spherical and cylindrical mirrors, alone and in combination.

## II. SPHERICAL MIRROR—MERIDIAN RAYS

When the plane of curvature of the section from which reflection takes place coincides with the plane of incidence, the rays are called meridian rays and the plane of curvature may be called the meridian plane.

In Figure 1 which illustrates meridian rays  $ON$  is a meridian section of a concave spherical mirror of radius  $R$  cm.  $C$  is the centre of curvature.  $P'O$  is an incident ray with grazing angle of incidence  $i$ ,  $i$  being the angle between  $P'O$  and the tangent through  $O$ .  $OQ'Q$  is the corresponding reflected ray. If the aperture is small an incident ray  $P'N$  is reflected along  $NQ'$ .

Consider a point object  $P$  at a distance  $p$  from  $O$ . Then  $Q$  will be the image of  $P$  where the angle  $PNP'$  equals the angle  $QNQ'$ . Let the angles at  $P'$ ,  $Q'$ , and  $C$  be  $\alpha$  (all equal for small aperture) and the angles at  $P$  and  $Q$  be  $\epsilon$  and  $\delta$  respectively. Let  $p$  and  $q$  be distances of object  $P$  and image  $Q$  from the pole of the mirror. A relation between  $p$  and  $q$  for small aperture may be established as follows:

$$P'O = OQ' = R \sin i.$$

\* Department of Physics, University of Western Australia.

† KIRKPATRICK, P., and BAEZ, A. V.—*J. Opt. Soc. Amer.* **38**: 766 (1948).

Assume  $PN=p$  and  $\alpha$  and  $\epsilon$  small. Therefore

$$p\epsilon=\alpha R \sin i.$$

Now assume  $Q'N=R \sin i$ . Therefore

$$q\delta=\alpha R \sin i.$$

Since  $\delta=2\alpha-\epsilon$  these two expressions yield

$$1/p+1/q=2/R \sin i=1/f_m,$$

where  $f_m$  is the meridian focal length and equal to  $\frac{1}{2}R \sin i$ .

The above two assumptions may be too inaccurate when  $i$  is small (unless the aperture is vanishingly small). It is interesting to examine the case of  $i$  small as well as  $\epsilon$  and  $\alpha$ .

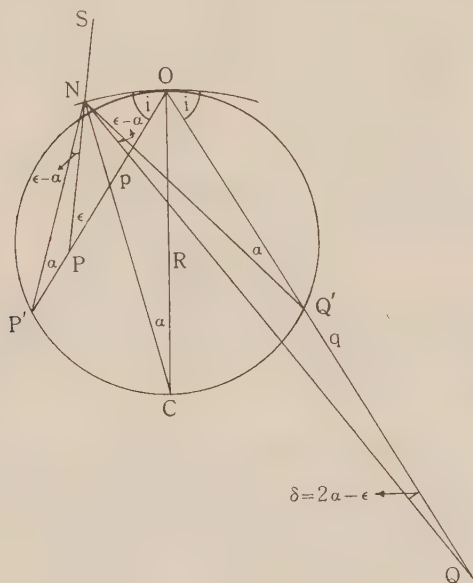


Fig. 1

In Figure 1 imagine the chord  $NO$  drawn. Then in the triangle  $PNO$

$$p/\sin PNO=\text{chord } NO/\sin \epsilon=(2R \sin \frac{1}{2}\alpha)/\sin \epsilon.$$

Now

$$\begin{aligned} \text{the angle } PNO &= \pi - SNO \\ &= \pi - (\epsilon + NOP) \\ &= \pi - (\epsilon + i - \frac{1}{2}\alpha). \end{aligned}$$

Therefore

$$p/\sin (\epsilon + i - \frac{1}{2}\alpha)=(2R \sin \frac{1}{2}\alpha)/\sin \epsilon.$$

Hence

$$p \sin \epsilon=2R \sin \frac{1}{2}\alpha \sin (\epsilon + i - \frac{1}{2}\alpha). \quad \dots \dots \dots (1)$$

Similarly in the triangle  $QNO$  we get

$$q/\sin QNO=\text{chord } NO/\sin \delta=(2R \sin \frac{1}{2}\alpha)/\sin \delta.$$

Now the angle  $QNO=(i+\frac{1}{2}\alpha-\delta)$ .

Hence

$$q \sin \delta=2R \sin \frac{1}{2}\alpha \sin (i+\frac{1}{2}\alpha-\delta). \quad \dots \dots \dots (2)$$



(From (1) and (2) we get as before, when  $i$  is not small,  $p\varepsilon = R\alpha \sin i = q\delta$ .)  
All the angles now being small (1) and (2) yield

$$p\varepsilon = R\alpha(\varepsilon + i - \frac{1}{2}\alpha), \quad \dots\dots\dots (1')$$

$$q\delta = R\alpha(i + \frac{1}{2}\alpha - \delta). \quad \dots\dots\dots (2')$$

Therefore

$$p/R\alpha - 1 = (2i - \alpha)/2\varepsilon,$$

and

$$q/R\alpha + 1 = (2i + \alpha)/2\delta.$$

Therefore

$$2\varepsilon = (2i - \alpha)/(p/R\alpha - 1),$$

and

$$2\delta = (2i + \alpha)/(q/R\alpha + 1).$$

Adding we get, since  $\delta + \varepsilon = 2\alpha$ ,

$$4\alpha = (2i - \alpha)/(p/R\alpha - 1) + (2i + \alpha)/(q/R\alpha + 1).$$

This reduces finally to

$$\frac{1}{q + R\alpha} + \frac{1}{p - R\alpha} \cdot \frac{2 - \alpha/i}{2 + \alpha/i} = \frac{1}{(1 + \alpha/2i)iR/2}.$$

Now substitute  $f_m = Ri/2$  and  $a = \alpha/i$ :

$$\frac{1}{q + 2af_m} + \frac{1}{p - 2af_m} \cdot \frac{2 - a}{2 + a} = \frac{1}{f_m(1 + a/2)}. \quad \dots\dots\dots (3)$$

This is identically equation (3) of Kirkpatrick and Baez.

When  $\alpha \ll i$  this equation reduces to the form

$$1/p + 1/q = 2/Ri.$$

In other words, in the case of meridian rays the formula  $1/p + 1/q = 2/R \sin i$  applies for all values of  $i$  (including  $\pi/2$ ) provided  $\varepsilon$  and  $\alpha$  are always small and provided that when  $i$  is also small,  $\alpha \ll i$ . If the latter condition does not hold equation (3) applies.

### III. SPHERICAL MIRROR—SAGITTAL RAYS

When the plane of curvature of the section from which reflection takes place is perpendicular to the plane of incidence, the rays are called sagittal rays and the plane of curvature may be called the sagittal plane.

In Figure 2 which illustrates sagittal rays and a sagittal section ( $ON$ ) of the mirror, the symbols have precisely the same significance as in Figure 1. In this case  $Q'$  is accurately the conjugate of  $P'$  where  $P'CQ'$  is perpendicular to the plane  $CON$ .  $Q$  is located as follows:  $P'NQ'C$  is the plane of incidence for the ray  $P'N$ . Rotate the plane about  $NC$  until it passes through  $P$ . Then if it intercepts  $OQ'$  in  $Q$ ,  $NQ$  will be the reflected ray corresponding to the incident ray  $PN$  and the angle  $Q'NQ$  equals the angle  $PNP'$ .

To obtain a relation between  $p$  and  $q$ :

$$P'O = OQ' = R/\sin i.$$

Assume  $PN = p$  and  $\alpha$  and  $\varepsilon$  small. Therefore

$$p\varepsilon = \alpha R/\sin i.$$

Now assume  $Q'N = R/\sin i$ . Therefore

$$q\delta = \alpha R/\sin i.$$

Since  $\delta = 2\alpha - \varepsilon$  these two expressions yield

$$1/p + 1/q = (2 \sin i)/R = 1/f_s,$$

where  $f_s$  is the sagittal focal length and equal to  $R/2 \sin i$ .

As before it is interesting to investigate the case of  $i$  small as well as  $\alpha$  and  $\varepsilon$ .

In Figure 2 imagine the chord  $NO$  drawn. Then in the triangle  $PNO$

$$p/\sin PNO = \text{chord } NO/\sin \varepsilon = (2R \sin \frac{1}{2}\alpha)/\sin \varepsilon \sin i.$$

Now

$$\text{the angle } PNO = \text{angle } P'NO - (\varepsilon - \alpha),$$

$$= (\frac{1}{2}\pi - \frac{1}{2}\alpha) - (\varepsilon - \alpha),$$

$$= \frac{1}{2}\pi - (\varepsilon - \frac{1}{2}\alpha).$$

Therefore  $\sin PNO = \cos (\varepsilon - \frac{1}{2}\alpha)$ . Substituting in the previous equation we get

$$p/\cos (\varepsilon - \frac{1}{2}\alpha) = (2R \sin \frac{1}{2}\alpha)/\sin \varepsilon \sin i.$$

Therefore

$$p \sin \varepsilon = 2 \sin \frac{1}{2}\alpha \cos (\varepsilon - \frac{1}{2}\alpha) R/\sin i.$$

Similarly

$$q \sin \delta = 2 \sin \frac{1}{2}\alpha \cos (\delta - \frac{1}{2}\alpha) R/\sin i.$$

(From these two equations we get, when  $i$  is not small, as before  $p\varepsilon = \alpha R/\sin i = q\delta$ .)

When all the angles are small the two equations yield

$$p\varepsilon = \alpha R/i = q\delta,$$

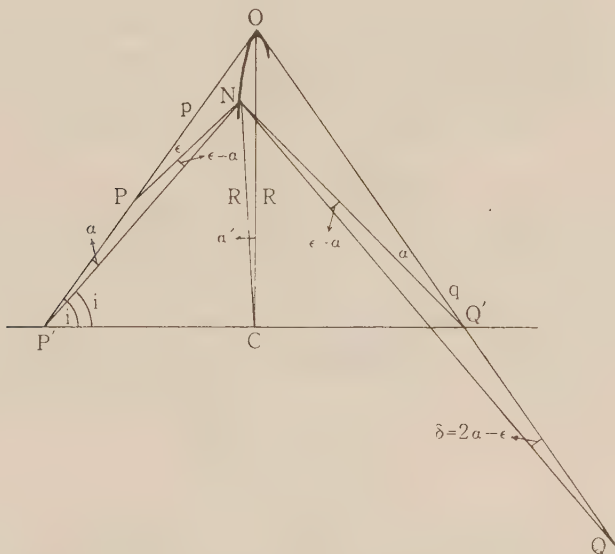


Fig. 2

and finally

$$1/p + 1/q = 2i/R.$$

In other words, in the case of sagittal rays the formula  $1/p + 1/q = (2 \sin i)/R$  applies for all values of  $i$  (including  $\frac{1}{2}\pi$ ) provided  $\varepsilon$  and  $\alpha$  are always small. The further restriction on  $\alpha$  when  $i$  is small (necessary in the case of meridian rays) is not necessary in this case.

It is interesting to note that in the case of sagittal rays, the condition that  $\alpha$  be small does not necessarily restrict the aperture  $\alpha'$ . For since  $\sin \frac{1}{2}\alpha' = (\text{chord } ON)/2R$  and since  $\sin \frac{1}{2}\alpha = (\text{chord } ON) (\sin i)/2R$ , we have

$$\sin \frac{1}{2}\alpha = \frac{1}{2}\alpha = (\sin \frac{1}{2}\alpha') \sin i.$$

Consequently if  $i$  is small  $\alpha'$  may be large. Only when  $i$  is large must the aperture be restricted. In the sagittal case therefore much larger apertures are in general permitted than in the meridian case.

## IV. ROWLAND LOCUS AND FOCAL LOCUS

In Figure 3,  $PO$  is the direction of any ray incident on a spherical mirror at  $O$ , and  $OQ'_s$  is the direction of the corresponding reflected ray.

$$OP'_m = R \sin i = OQ'_m,$$

and

$$OF'_m = (R \sin i)/2 = OF_m.$$

Assume the aperture small enough for the condition  $\alpha < i$  to hold when  $i$  is small. Then for all angles of incidence if the object is at  $P'_m$  the meridian image will appear at  $Q'_m$  with unit magnification. The circle  $OC$  (diameter  $R$ ) may therefore be termed the meridian Rowland locus.

If the object  $P$  is at infinity the meridian image will appear at  $F'_m$ . The principal focus for meridian rays will therefore always lie on the smaller circle (diameter  $\frac{1}{2}R$ ). This circle may therefore be termed the meridian focal locus.

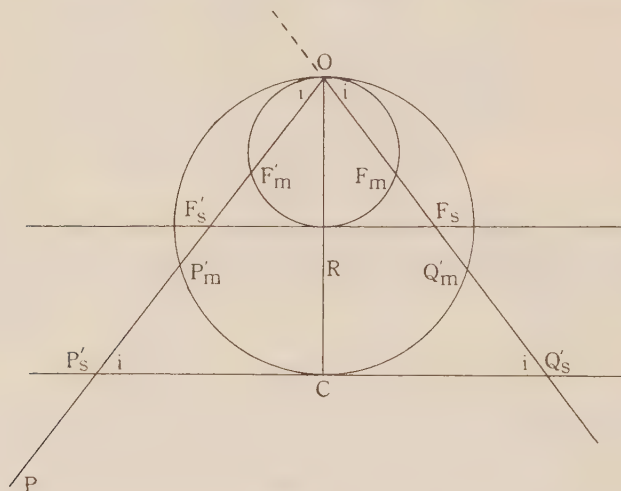


Fig. 3

It may be noted that if the object lies between  $F'_m$  and  $O$  the meridian image will be virtual and will lie somewhere between  $O$  and infinity along the line  $Q'_sO$  produced.

The case for sagittal rays is illustrated in the same figure.

$$OP'_s = R/\sin i = OQ'_s,$$

and

$$OF'_s = R/2 \sin i = OF_s.$$

For all angles of incidence if the object is at  $P'_s$  the sagittal image will appear at  $Q'_s$  with unit magnification. The axis  $P'_sCQ'_s$  extended will therefore be the sagittal Rowland locus.

If the object  $P$  is at infinity the image will appear at  $F'_s$ . The principal focus for sagittal rays will therefore always lie on the line  $F'_sF_s$  (extended) and this line (bisecting  $OC$ ) is the sagittal focal locus. It may be noted that if the object lies between  $F'_s$  and  $O$  the sagittal image will be virtual and will lie somewhere between  $O$  and infinity along the line  $Q'_sO$  produced.

In the case of a cylindrical mirror in such an orientation that the sagittal radius of curvature  $R_s$  is infinite, the sagittal focal and Rowland loci are at infinity. This means that the sagittal rays form a virtual image in the same position as that formed by a plane reflector. The meridian rays behave as the meridian rays of a spherical mirror of the same radius of curvature (say  $R_m$ ). In the other orientation ( $R_m$  infinite) the meridian focal and Rowland loci are at infinity.

### V. THEORY OF "CROSSED" MIRRORS

It is apparent from the foregoing discussion that a concave spherical mirror may be regarded as forming two distinct astigmatic images. Since the grazing angle of incidence  $i$  is the same for meridian and sagittal rays, one image is formed by the meridian rays for which the focal length is  $(R \sin i)/2$  and the other by the sagittal rays for which the focal length is  $R/2 \sin i$ . The former is astigmatic perpendicular to the meridian plane and the latter parallel to the meridian plane. The images coincide and the astigmatism corrected only when  $i$  is  $\frac{1}{2}\pi$ . In the case of a cylindrical mirror in such an orientation that the sagittal radius  $R_s$  is infinite, the images never coincide and astigmatism is always present.

The astigmatism will in general be corrected and the final image (or images) superposed if two "crossed" mirrors are used and placed at a given distance apart. Mirrors are said to be "crossed" when meridian rays from the first become sagittal rays for the second and vice versa. It remains to determine the requisite distance apart of the mirrors for superposition of the final images.

For the first mirror (object distance  $p$  and focal lengths  $f_m$  and  $f_s$ ) the image distances  $q_m$  and  $q_s$  are given by

$$1/f_m = 1/q_m + 1/p,$$

and

$$1/f_s = 1/q_s + 1/p.$$

That is

$$q_m = \frac{pf_m}{p - f_m},$$

$$q_s = \frac{pf_s}{p - f_s}.$$

Let  $d$  be the distance apart of the mirrors. Then for the rays constituting the meridian rays for the second mirror the object distance for the mirror is

$$p_m = d - q_s.$$

For the sagittal rays the object distance is

$$p_s = d - q_m.$$

We have therefore for the second mirror

$$1/f'_m = 1/q'_m + 1/(d - q_s),$$

and

$$1/f'_s = 1/q'_s + 1/(d - q_m).$$

Whence

$$q'_m = \frac{f'_m(d - q_s)}{(d - q_s - f'_m)}, \text{ and } q'_s = \frac{f'_s(d - q_m)}{(d - q_m - f'_s)}.$$



For superposition of final images  $q'_m = q'_s = q$ . Equating  $q'_m$  and  $q'_s$  and substituting for  $q_m$  and  $q_s$ , we get

$$\frac{f'_m \left( d - \frac{pf_s}{p-f_s} \right)}{d - \frac{pf_s}{p-f_s} - f'_m} = \frac{f'_s \left( d - \frac{pf_m}{p-f_m} \right)}{d - \frac{pf_m}{p-f_m} - f'_s}.$$

Rearranging terms we get

$$d^2 - d(q_m + q_s) + \frac{p^2 f_m f'_m (f_s - f'_s) - p^2 f_s f'_s (f_m - f'_m)}{(p - f_m)(p - f_s)(f'_m - f'_s)} = 0.$$

So far this is quite general. Assume  $f_m = f'_m$  and  $f_s = f'_s$ . In the case of spherical mirrors this means that

$$R \sin i = R' \sin i', \text{ and } R/\sin i = R'/\sin i'.$$

In the case of cylindrical mirrors it means

$$R_s = R'_s = \infty, \text{ and } R_m \sin i = R'_m \sin i'.$$

(The other possible "crossed" position of two cylindrical mirrors would give  $f_s = f'_m$ .) If  $i$  is the same for each mirror the two mirrors in each case are identical. With the above assumptions the previous equation becomes

$$d^2 - d(q_m + q_s) = 0,$$

giving

$$d = 0 \text{ or } d = q_m + q_s.$$

When  $d = 0$ , in the case of spherical mirrors the formulae for the combination become

$$1/F = 1/f_m + 1/f_s = 1/q + 1/p,$$

and a single final image of magnification  $-q/p$  appears. This applies when  $f_s > f_m$  (which includes cylindrical mirrors).

When  $d = q_m + q_s$ , in the case of spherical mirrors the formulae for the second mirror become

$$1/f_m = 1/q + 1/q_m,$$

and

$$1/f_s = 1/q + 1/q_s.$$

Comparing these with the formulae for the first mirror we see that  $p = q$ . The final magnification for the meridian-sagittal case is  $q_m/q_s$ . Whereas for the sagittal-meridian case it is  $q_s/q_m$ . The two superposed images therefore are of reciprocal or conjugate magnifications. This applies when  $f_s > f_m$  (which includes cylindrical mirrors).

When  $q_m = -q_s$  we have one image. This is the particular case of  $d = 0$  for which  $p = q$ , and magnification  $= -1$ .  $p$  is then equal to  $2f_m f_s / (f_m + f_s)$ , or  $2f_m$  in the case that  $f_s > f_m$ . It may be shown that this is the upper limiting value of  $p$  for positive (and therefore practicable) values of  $d$ .

A single final image is produced with crossed mirrors only if  $d = 0$ , or if  $d$  is small compared with  $p$  and  $q$ . The minimum value of  $d$  is the linear aperture of either of the mirrors (if the dimensions of each are the same). It is always possible in a given case to find the appropriate value of  $p$  for this  $d$  and obtain the two images in focus. If one is of large magnification the presence of the other may not be serious. However, only half the reflected energy contributes to the wanted image.

Another possible solution, theoretically, is to use a single mirror whose meridian radius of curvature is  $R_m$  and whose sagittal radius of curvature is  $R_s$ , where

$$R_s/R_m = \sin^2 i.$$

This would make  $f_m = f_s$ . The case is illustrated in Figure 4 where the dotted lines represent the Rowland loci. From geometry it is apparent that for an object anywhere along  $OP$ ,  $F$  is the principal focus and  $OP$  is the object distance for unit magnification.  $C_m$  is the centre of curvature for the meridian section and  $C_s$  for the sagittal section.

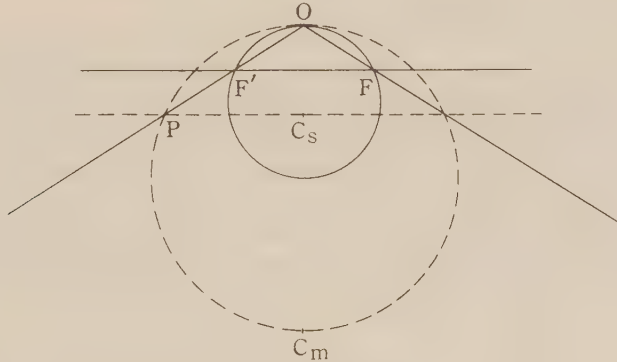


Fig. 4

Such a single mirror has the advantage of combining in one piece an optical system that will form unstigmatic achromatic images, and at the same time deviate the beam through any fixed angle for which the mirror is designed.

It is interesting to examine how closely two cylindrical mirrors would or would not achieve the same result. This will be deduced from the general case of any two mirrors being used in a "parallel" position.

## VI. THEORY OF "PARALLEL" MIRRORS

In this case meridian rays from the first mirror constitute meridian rays for the second mirror. Similarly to the previous case of "crossed" mirrors we get for the first mirror

$$q_m = \frac{pf_m}{p - f_m}, \text{ and } q_s = \frac{pf_s}{p - f_s}.$$

For the second mirror we now have

$$p_m = d - q_m \text{ and } p_s = d - q_s,$$

giving

$$1/f'_m = 1/q'_m + 1/(d - q_m),$$

and

$$1/f'_s = 1/q'_s + 1/(d - q_s).$$

For superposition of final images we get  $q'_m = q'_s = q$ , and therefore  $q'_m - q'_s = 0$ , yielding

$$f'_m(d - q_m)(d - q_s - f'_s) - f'_s(d - q_s)(d - q_m - f'_m) = 0.$$

This finally gives

$$d^2 - d(q_m + q_s) + \frac{p^2 f'_m f_s (f_m - f'_s) + p^2 f_m f'_s (f'_m - f_s)}{(p - f_m)(p - f_s)(f'_m - f'_s)} = 0.$$

So far this is quite general. In the case of non-spherical mirrors ( $R_m \neq R_s$ ) there are two possible "parallel" positions. The interesting one is that for which  $f_s = f'_m$  ( $= \infty$  for cylindrical mirrors). With the further condition that  $f_m = f'_s$  we again get

$$d^2 - d(q_m + q_s) = 0,$$

giving either  $d=0$  and a single final magnification  $-q/p$ , or  $d=q_m + q_s$  and two images of conjugate magnifications.

Now  $f_m = (R_m \sin i)/2$  and  $f'_s = R'_s/2 \sin i'$ . If  $i=i'$ ,  $R'_s/R_m = \sin^2 i$ . In the case of cylindrical mirrors the sagittal rays of the first mirror and meridian rays of the second are reflected as from plane surfaces and so these planes, as they contribute nothing geometrically to the image-forming property of the combination, may be absent. Consequently a single surface with radii of curvature, in sections at right angles, in the ratio of  $\sin^2 i$  automatically ensures  $i=i'$ , and will be equivalent to two cylindrical mirrors mounted as described above with  $d$  zero. We are then reduced to the single mirror previously discussed.

# OBSERVATIONS OF THE SPECTRUM OF HIGH-INTENSITY SOLAR RADIATION AT METRE WAVELENGTHS

## III. ISOLATED BURSTS

By J. P. WILD\*

[Manuscript received June 22, 1950]

### Summary

Observations are described of the spectrum of "isolated bursts" of solar radio-frequency radiation in the frequency range 70–130 Mc/s. These bursts last for a few seconds and have a bandwidth of the order of tens of megacycles per second. Prior observations indicate that they are not circularly polarized. They occur sporadically, often in small groups; many hours sometimes elapse between successive bursts or groups. Although, in general, their spectra show diverse features, some of them (referred to as "type III" bursts) are of a distinct type characterized by a rapid drift, with time, of the frequency of maximum intensity towards the lower frequencies, at a rate of the order of 20 Mc/s. per second. Characteristics of the spectra of type III bursts are described in detail.

The results are discussed and hypotheses of origin examined. It is shown in particular that the frequency drift of type III bursts cannot be attributed to the selective group retardation of waves in the solar atmosphere emanating from a fixed source. The frequency drift may, however, be associated with the rapid motion of a source travelling outwards through the solar atmosphere.

## I. INTRODUCTION

This is Part III of a series describing some observations of the radio spectrum of high-intensity solar radiation in the frequency range 70–130 Mc/s. In Part I(1), the apparatus was described and means of distinguishing between classes of burst according to their spectra were given. It was shown, in particular, that of the broad group of bursts that occur at *sporadic* intervals (apparently irrespective of the presence or absence of a general background level of "enhanced" radiation), a distinct class of burst, associated with solar flares, could be recognized from spectral characteristics; these bursts, referred to as "outbursts", were treated in Part II(2). The present paper deals with the remaining sporadic-type bursts, referred to as *isolated bursts*.† Observations of these bursts with single-frequency equipment have previously been described by Pawsey(3) and Payne-Scott(4), and they have been found to have the following characteristics:

- (i) They are typically of short duration (some seconds).
- (ii) They often occur nearly simultaneously at widely different frequencies.
- (iii) They show no circular polarization.

\* Division of Radiophysics, C.S.I.R.O., University Grounds, Sydney.

† The term was introduced by Pawsey(3). The bursts are referred to by Payne-Scott(4) under the category of "unpolarized bursts", which include "outbursts".



## II. DESCRIPTION AND CHARACTERISTICS OF OBSERVED SPECTRA

In the course of observations taken during 264 hours at intermittent times between February and June 1949, spectra of 32 isolated bursts, or small groups of them, were recorded. Each of the records was converted into time-frequency-intensity diagrams in the manner described in Part I. Some examples are shown in the accompanying coloured plates. Inspection of the diagrams indicates at once that the spectra differ widely. An attempt is made in this section to describe the observed spectra and indicate possible characteristic features.

### (a) *Frequency and Time Profiles*

The variation of intensity with frequency at a particular time will be referred to as a *frequency profile*, and the variation of intensity with time at a particular frequency as a *time profile*. These profiles are cross sections of the contour diagrams along horizontal and vertical lines respectively.

Frequency profiles of the isolated bursts usually had the form of fairly smooth curves with no sudden variations within frequency intervals of the order of a few megacycles per second. (In this respect the isolated bursts differed from "outbursts" and "storm bursts" discussed elsewhere in this series of papers.) The bandwidth of the frequency profiles, measured between points of quarter-maximum intensity, was rarely less than 10–15 Mc/s. and often much greater: in many cases the bandwidth was too broad to allow measurement in the restricted frequency range under observation. In simple cases (see, for instance, Plate 2a), the frequency profiles showed only one maximum; such profiles were in general unsymmetrical, and the variation of intensity with frequency on the high frequency side of the maximum tended to be more gradual than that on the low frequency side. In other cases (e.g. Plate 1b) the frequency profiles showed more than one maximum.

Three exceptional bursts were observed in which frequency profiles during part of the burst showed sudden intensity variations. These bursts are illustrated in Plate 3. In each case an abrupt cut-off in frequency appears on one side of a high selective peak. In two cases the cut-off lies on the low frequency side of the peak, and in one case on the high frequency side.

The *time* profiles of simple isolated bursts showed a gradual rise to a maximum value, followed by a gradual decay to the original level. Those of more complex isolated bursts (e.g. Plate 4a–c) had several maxima. In a period of 1 second the intensity did not commonly change by a factor of more than four, although this rate was considerably exceeded in a few cases. The time of rise tended to be less than the time of decay; on the average about 0.6 of the total lifetime was spent in decay.

On occasions when the intensity was observed to decrease by a large factor (say,  $2^4$ ), the points at which the time profiles crossed the contours tended to be approximately equally spaced. This is in agreement with results of Williams(5) and Payne-Scott(4) who found that the intensity of bursts at any one frequency tended to decay exponentially with time.

*(b) Temporal Variations of Frequency Profiles: Type III Bursts*

In bursts with simple frequency profiles that reach a single maximum value, the frequency profiles usually changed with time in one of two ways: the frequency of maximum intensity either (i) remained approximately constant throughout the lifetime of the burst (e.g., Plate 1*a*), or (ii) drifted rapidly with time (e.g., Plate 2*a*). Bursts of the latter kind appeared to conform quite remarkably to a pattern: the drift was almost invariably in the direction of decreasing frequency, and the rate of drift seldom differed from an average value of about 20 Mc/s. per second by a factor of more than two. Such bursts were observed repeatedly, and it seems reasonable to regard them as a specific spectral type. In accordance with Part I, in which three distinct types of burst were indicated, those that show this rapid frequency drift will be designated as "type III" bursts. They will be described in detail in a later section. The occurrence of type III bursts is consistent with the results of observers of this Laboratory (6, 4) who worked with single-frequency equipment; they found that when bursts appeared almost simultaneously at two different frequencies, that at the higher frequency normally started slightly in advance of that at the lower frequency.

*(c) Groups of Bursts and Complex Bursts*

Although isolated bursts sometimes occurred as single disturbances preceded and followed by undisturbed periods of several hours' duration, they were frequently found to occur in small groups, two or three bursts arriving in quick succession. Without a rigorous statistical investigation it is not possible to state definitely that this "bunching" effect is not due to chance coincidences in a random distribution, but the observations suggested that the bunching effect was considerably greater than that to be expected by chance. This effect suggests that bursts with complex spectra (i.e. complex as regards both their frequency and time variations) may properly be regarded as a number of simple bursts clustered together in the time-frequency plane. A second factor supporting this interpretation is that parts of complex bursts often resembled typical kinds of simple bursts. For instance, the complex bursts shown in Plate 4 each appear to contain a typical "type III" burst (cf. Plate 2*a*).

Bursts that occurred within a short period (say, a minute) of one another—whether different bursts of a group, or different components of a complex burst—were sometimes found to have markedly similar spectra. Such bursts tended to occur in descending order of amplitude. Examples of this "repetition" effect are to be found in Plates 2*c* and 4*a*.

*(d) Energy Spectra*

To study the variation of the amplitude of a burst with frequency in a way which is independent of time, we can evaluate the total energy per unit frequency interval that flows normally through unit area at the earth during the burst. This is given by

$$E_f = \int S_f dt,$$

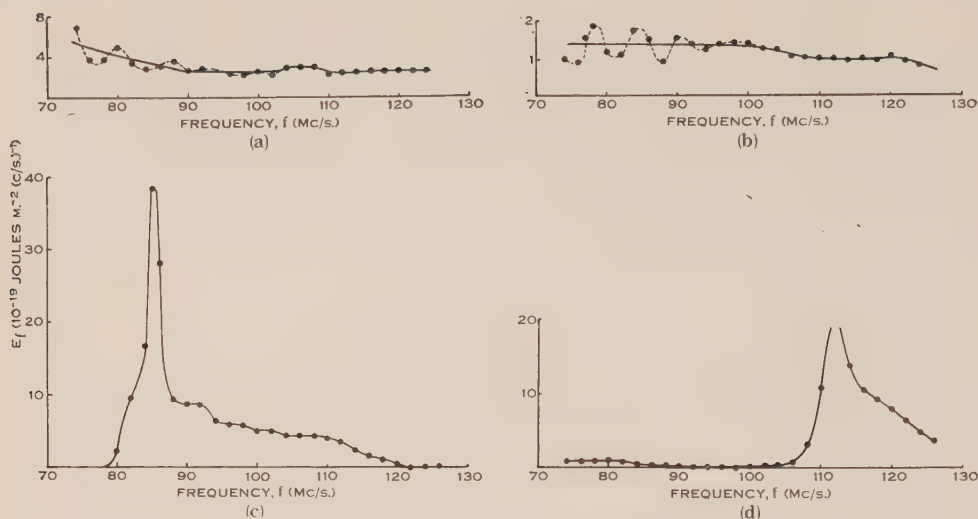


Fig. 1.—“Energy” spectra of four isolated bursts.

(a) and (b) correspond to Plate 2b and 2a respectively (type III).

(c) and (d) correspond to Plate 3a and 3b respectively (note “cut-off” and peak in each case).

Note.—In (a) and (b), the fluctuations at the lower frequencies are believed to be instrumental in origin (see Part I).

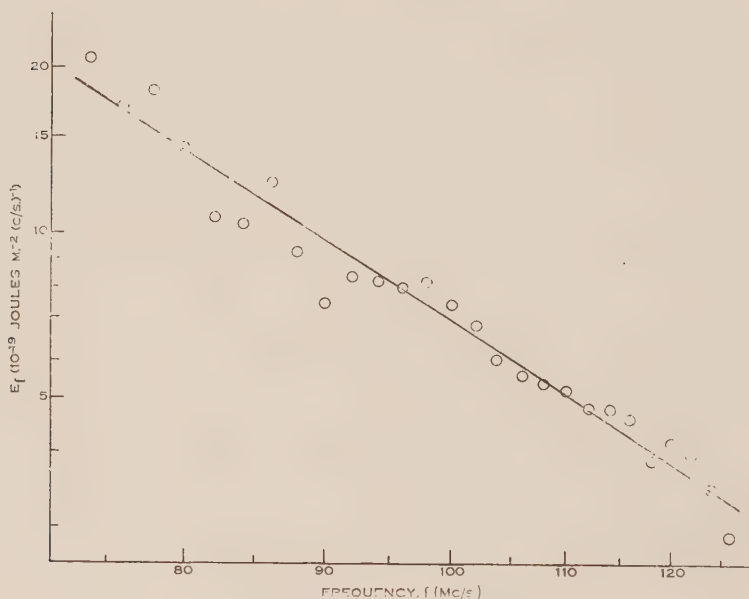


Fig. 2.—A statistical spectrum giving the average distribution of energy with frequency of observed isolated bursts. ( $E_f$  represents the *total* energy per unit frequency interval flowing normally through unit area at the earth during 29 isolated bursts.)  $E_f$  and  $f$  are both plotted on a logarithmic scale. The line drawn through the observational points corresponds to the law

$$E_f \propto f^{-3.25}.$$

where  $S_f$  is the "intensity", as plotted in the diagrams,\* and  $f$  and  $t$  denote frequency and time respectively. The "energy spectrum", defined as the variation of  $E_f$  with  $f$ , of a burst can be determined approximately from the contour diagrams by graphical integration. Four examples are given in Figure 1 (a-d): (a) and (b) refer to type III bursts and show only a gradual variation of  $E_f$  with  $f$ ; (c) and (d) refer to bursts that exhibit an abrupt low-frequency cut-off, and each shows a selective peak with a gradually declining "tail" on the high-frequency side.

To obtain a measure of the *average* frequency-distribution of the energy of the observed isolated bursts, we can add all the energy spectra of the separate bursts. The result of adding the integrated spectra of 29 of the 32 observed burst groups is shown in Figure 2. (In order to ensure that any one burst did not have a large effect on the final result, those bursts which contributed more than 20 per cent. of the total energy in any part of the frequency range were excluded from the summation. Three bursts were thus excluded.) The curve shows a marked decline as the frequency increases. If an attempt is made to represent the curve by a single power function  $f^{-n}$ , a best fit is obtained by choosing a value of  $n$  of about 3.25; plausible values of  $n$  lie between 2.8 and 3.6.

### III. TYPE III BURSTS

We now consider in more detail the type III bursts, which are characterized by a rapid drift with time of the frequency of maximum intensity towards the lower frequencies. Some examples of the type are given in Plate 2. About half the groups of isolated bursts recorded were found to contain at least one type III burst. Some 20 type III bursts were recorded in all, but only about half were sufficiently well defined for useful analysis. Characteristics of these bursts are described below.

#### (a) *Frequency Drift*

The contour diagram of an idealized type III burst is sketched in Figure 3. The rate of change of frequency with time along the "ridge line", i.e. the reciprocal gradient of the ridge line in the figure, will be referred to as the *drift rate*. In all well-defined cases the drift rate became slightly slower as the frequency decreased, i.e. in terms of Figure 3, the ridge line approximated to a gentle curve with its centre of curvature below the line. As will be shown in Section IV (b), the nature of this curvature is such that the negative drift rate is approximately proportional to the frequency, i.e.

$$-\frac{df}{dt} = \frac{f}{\alpha}, \quad \dots \dots \dots (1)$$

where  $t$  and  $f$  specify the time and frequency of points on the ridge line and  $\alpha$  is constant (of the order of a few seconds) for a given burst. Approximate values of  $\alpha$  for observed type III bursts are given in Table 1.

\* The quantity  $S_f df$ , defined in Part I, is the accepted *plane-polarized component* of the power flowing normally through unit area at the aerial in the frequency interval between  $f$  and  $f+df$ . Thus the quantity  $E_f$  also refers to only one plane of polarization.



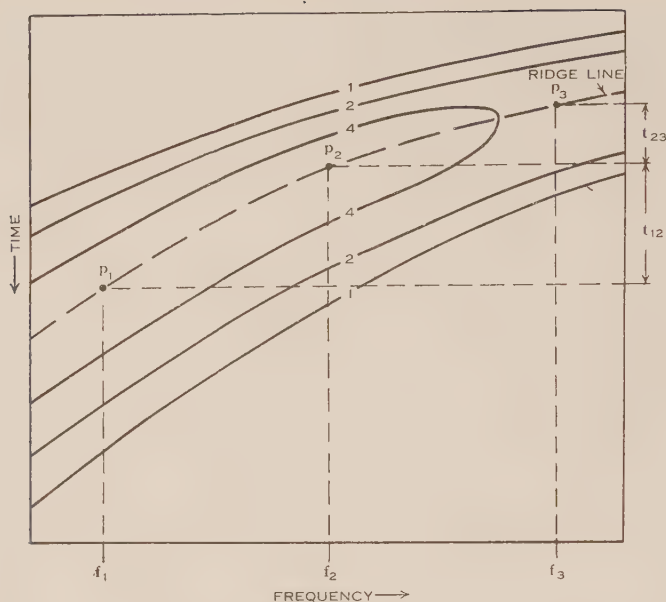


Fig. 3.—Contour diagram of an idealized type III burst. The numbers on the contours denote intensity in arbitrary units.

TABLE I

APPROXIMATE VALUES OF  $\kappa$  AND  $\alpha$  FOR 10 OBSERVED TYPE III BURSTS

Time of Occurrence of Burst (G.M.T.)	Illustrated in Plate	$\kappa \times 10^8$	$\alpha$ (sec.)
18.iii.49; 05 hr. 13 min.	2 <i>d</i>	0.64	5.8
21.iii.49; 02 hr. 43 min.	2 <i>b</i>	0.96	4.8
4.iv.49; 03 hr. 23 min.	—	0.76	4.1
6.iv.49; 01 hr. 21 min.	—	1.5	1.9
6.iv.49; 02 hr. 12 min.	—	1.2	3.1
11.iv.49; 05 hr. 59 min.	4 <i>b</i> (part)	0.90	4.7
30.v. 49; 01 hr. 13 min.	4 <i>c</i> (part)	0.90	3.0
30.v. 49; 01 hr. 20 min.	—	0.62	7.2
31.v. 49; 00 hr. 38 min.	2 <i>c</i> (part)	0.94	4.8
31.v. 49; 01 hr. 09 min.	2 <i>a</i>	1.35	6.2

### (b) Lifetime and Bandwidth

The lifetime at any one frequency tended to increase slightly with decreasing frequency. Measurement of the mean lifetime was possible in 10 cases; in all but one of these cases the values (measured between points of quarter-maximum intensity) lay between 2.9 and 3.8 seconds. On the average, slightly more than half the lifetime was spent in decay. In the majority of type III

bursts the time profiles appeared to be nearly symmetrical in shape (e.g. Plate 2*a*), but in a few cases (e.g. Plate 2*b*) the decay portion was markedly greater than the rise portion.

The measurement of the bandwidth of these bursts was generally limited by the narrowness of the observed frequency range, many of the bandwidths being greater than 60 Mc/s. The lowest bandwidth measured was about 50 Mc/s. However, the bandwidth of type III bursts may be regarded as a dependent property determined by the drift rate and the lifetimes measured over a sufficiently broad frequency range.

### (c) Decay Constant

As already stated, the decay of time profiles of observed isolated bursts appeared to be approximately exponential in form, i.e. the intensity of the decay portion could be represented approximately by a curve proportional to  $e^{-kt}$ . We now consider the variation of the "decay constant",  $k$ , with frequency for the observed type III bursts.

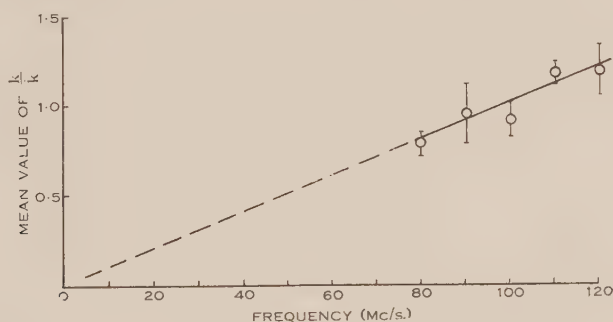


Fig. 4.—The variation with frequency of the average value of the normalized decay constant,  $k/\bar{k}$ , for 12 observed type III bursts. ( $k$  is the constant in the decay factor  $e^{-kt}$  of an intensity curve at any one frequency, and  $\bar{k}$  is a mean value of  $k$  over the frequency band for a single burst.) Probable errors are indicated in the figure.

Of 12 type III bursts suitable for measurement, five exhibited a steady increase of  $k$  with frequency, two more *suggested* an increase of  $k$  with frequency, and the remaining five were inconclusive; in no instance was there a suggestion of a decrease of  $k$  with frequency. The statistical variation of decay constant with frequency is shown in Figure 4. (The ordinate in this figure is the average value, for the 12 bursts, of  $k/\bar{k}$  at a given frequency, where the normalizing factor  $\bar{k}$  is the mean value of  $k$  over the frequency range for a given burst.) The figure suggests that the decay constant in the observed frequency range may be approximately represented by

$$k = \alpha f, \quad \dots \dots \dots (2)$$

where  $\alpha$  is a constant (of the order of  $10^{-8}$ ) for a given burst. Approximate values of  $\alpha$  for observed bursts are given in Table 1.

## IV. DISCUSSION OF RESULTS

Little is known of the origin of isolated bursts. The two other general classes of burst discussed elsewhere in this series of papers provide us with some hint about their origin because each shows a definite correlation with solar phenomena capable of optical observation. But in the case of isolated bursts we have at present no evidence except that provided by radio-frequency observation. In the present discussion possible causes of the bursts will be examined in the light of the foregoing evidence of radio spectra.

*(a) Comparison of General Results with a Theoretical Model*

Because they are short-lived phenomena with relatively rapid temporal variations, isolated bursts are likely to originate in localized rather than distributed disturbances on the sun. Furthermore, considerations of the propagation and escape of radio waves in an ionized gas indicate that a source generating bursts at metre wavelengths would be situated in the more rarefied layers of the solar atmosphere, i.e. in the solar corona.

The effects of *localized* (one-point), *instantaneous* electrical disturbances in the solar corona have been investigated theoretically by Jaeger and Westfold(7). They found that such a disturbance may be capable of producing a short-lived radio burst observable at the earth, and that many predicted properties of such bursts are consistent with observed properties derived from single-frequency measurements. It is therefore of interest to test these predictions with the spectral evidence given in this paper.

The intensity-frequency characteristics of a Jaeger-Westfold burst are shown diagrammatically in Figure 5 (a). The spectrum shows a selective peak at the frequency  $f_0$ , falling abruptly to zero on the low-frequency side, and more gradually on the high-frequency side. Physically, the frequency  $f_0$  is the natural plasma frequency in the region of the source and may assume different values according to the electron density of the atmosphere at the source. If observed in a frequency range that contains  $f_0$  (e.g. the range  $R_1$  in the figure), the spectrum of a burst should show the characteristic peak, low-frequency cut-off and high-frequency "tail". Spectra of this kind have in fact been observed in two cases (see Plate 3a and b, and the corresponding "energy spectra" in Figure 1 (c) and (d)). If, on the other hand,  $f_0$  lies below the low-frequency limit of the observed range (e.g. the range  $R_2$  in the figure), then the spectrum should show a steady decrease of intensity with frequency. It follows that this kind of spectrum would be exhibited by all bursts except those which showed a characteristic low-frequency cut-off in the observed range. Inspection of the accompanying coloured plates indicates that this is certainly not true in all *individual* cases. However, the *average* spectrum (see Fig. 2) clearly shows the predicted trend. Thus the intensity-frequency relationships predicted by Jaeger and Westfold show some agreement with the observations.

The frequency-time characteristics of a Jaeger-Westfold burst are represented in Figure 5 (b). The *full line* in the figure represents the time of maximum intensity at a given frequency for the radiation emitted outwards from the source. Higher frequencies arrive before lower frequencies; this effect is due

to *selective group retardation*, i.e. to the variation of group velocity of waves of different frequency travelling through the ionized atmosphere. Thus if  $R_2$  represents the observed frequency range, the theory predicts a time-frequency spectrum which is qualitatively not unlike that of the characteristic type III bursts. However, it will be shown below in a quantitative analysis that the frequency drift of type III bursts is not consistent with that resulting from group retardation alone.

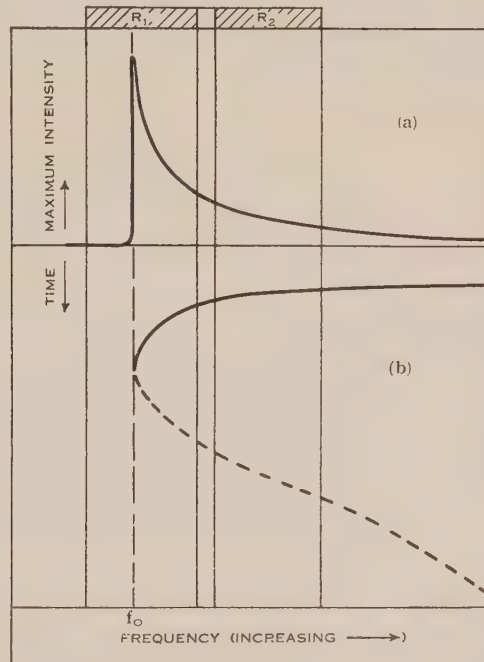


Fig. 5.—Characteristics of a Jaeger-Westfold burst :  
 (a) The variation of maximum intensity with frequency.  
 (b) The line represents relative times of arrival of different frequencies of maximum intensity (cf. "ridge line" in Fig. 3). The dotted portion refers to the "echo".

The theory further predicts that the main burst (as represented by the full line in the figure) will be followed by an "echo" reflected from the zero-refractive-index level of appropriate frequency. The time-frequency spectrum of this echo is represented by the dashed line in the figure. Jaeger and Westfold considered that the preponderance of "double-humped" bursts observed at single frequencies by Payne-Scott(4) supported their theory. Although, in the present observations, some instances suggestive of echoes were recorded (see, for instance, Plate 2b and the examples of "repetition" cited at the end of Section II (c)), none appeared to show the predicted time-frequency spectrum.



From this comparison of the spectral evidence with the theory of Jaeger and Westfold, it is concluded that:

- (i) The predicted intensity-frequency spectrum is in fair agreement with the observations.
- (ii) The predicted time-frequency spectrum does not appear to be in accordance with the observations.

(b) *On the Origin of Type III Bursts*

We now consider the type III burst in more detail, and examine, in particular, mechanisms that may be responsible for their characteristic frequency drift.

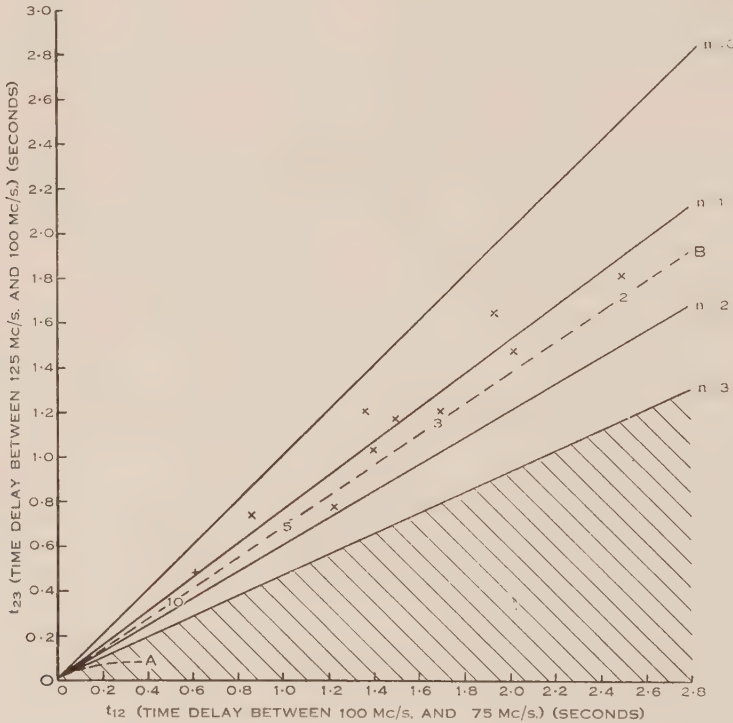


Fig. 6.—Diagram for studying the frequency drift of type III bursts. The coordinates,  $t_{12}$  and  $t_{23}$ , are defined by Figure 3. Experimental points are shown thus — x. The number  $n$  refers to the curvature law of the ridge line (drift rate  $\propto f^n$ ). Different hypotheses predict points in different regions of the diagram, as follows:

<i>Hypothesis</i>	<i>Region of Diagram</i>
Selective group retardation (Baumbach-Allen atmosphere)	Line OA
Selective group retardation (any atmosphere)	Shaded region
Outward-moving source of constant velocity (Baumbach-Allen atmosphere)	Line OB (The numbers on the line refer to velocities in $10^4$ km./sec.)

It is convenient to introduce a means of representing, as completely as possible, the frequency drift of a type III burst by a single point on a graph. Such a representation is shown in Figure 6 which is constructed in the following manner. Three points,  $P_1$ ,  $P_2$ , and  $P_3$ , are selected on the ridge line at frequencies  $f_1$ ,  $f_2$ , and  $f_3$  (see Fig. 3). The time interval between points  $P_1$  and  $P_2$  is denoted by  $t_{12}$  and that between  $P_2$  and  $P_3$  by  $t_{23}$ . Figure 6 then represents a plot of  $t_{12}$  against  $t_{23}$  when  $f_1=75$  Mc/s.,  $f_2=100$  Mc/s.,  $f_3=125$  Mc/s. A feature of the diagram is that it indicates not only the average "slope" of the ridge line in the two frequency ranges (proportional to the coordinates  $t_{12}$  and  $t_{23}$  respectively), but also the "curvature" of the ridge line. Thus if a burst showed a straight ridge line (i.e. constant drift rate), it would give a point on the  $45^\circ$ -line (marked " $n=0$ "). Similarly if the ridge line were curved and the drift rate given by

$$-\frac{df}{dt} = \frac{f^n}{\alpha},$$

points would lie on the straight line given by

$$\frac{t_{23}}{t_{12}} = \frac{f_2^{1-n} - f_3^{1-n}}{f_1^{1-n} - f_2^{1-n}}, \quad \dots\dots\dots (3)$$

the position of the point on the line being determined by the value of  $\alpha$ . Lines corresponding to  $n=1$ , 2, and 3 are drawn in the figure. The observational points are seen to lie between the lines  $n=0$  and  $n=2$ , the line  $n=1$  giving a fairly good fit. This is the basis for equation (1) in Section III. It will be shown in the subsequent discussion that different hypotheses regarding the cause of frequency drift predict points in different regions of the diagram. This enables us to test the hypotheses directly with the observations.

Possible mechanisms causing the frequency drift are :

- (i) Selective group retardation (as envisaged in the theory of Jaeger and Westfold).
- (ii) The outward motion, through the solar atmosphere, of a localized agency which excites plasma oscillations of continuously decreasing frequency. (A similar mechanism has been discussed in Part II with reference to "outbursts".)
- (iii) Some mechanisms in which the wave frequency is controlled by an external magnetic field, frequency drift being the result of a changing magnetic field.

The last of these three possibilities seems improbable in view of the observed absence of circular polarization in isolated bursts(3, 4). We shall therefore consider only the first two possibilities, and in these we shall neglect magnetic fields.

(i) *The Interpretation of Frequency Drift in Terms of Selective Group Retardation.*—The group velocity of radio waves in an ionized medium is given by

$$v = \mu c \quad (0 < \mu \leq 1), \quad \dots\dots\dots (4)$$

where  $\mu$  is the refractive index of the medium and  $c$  the free-space velocity of light. The waves can be propagated in regions where  $\mu > 0$ . Propagation in

the direction normal to surfaces of constant  $\mu$  is rectilinear. At oblique angles, however, propagation takes place along curved paths, but, apart from regions where the refractive index departs appreciably from unity, the paths approximate to straight lines (see Jaeger and Westfold 8). For this reason, and because of the added complexity introduced by curved paths, we shall at first consider only the case of rectilinear propagation.

In the absence of a magnetic field, the refractive index in the solar atmosphere is given by the Lorentz formula

$$\mu^2 = 1 - \frac{Ne^2}{\pi m f^2}, \dots\dots\dots (5)$$

where  $N$  is the electron density,  $e$  and  $m$  the electronic charge (e.s.u.) and mass, and  $f$  the frequency.

It follows from equation (4) that the interval between the times of arrival (at the observer) of two signals of frequency  $f_i$  and  $f_j$  originating simultaneously in a localized disturbance in the solar atmosphere is

$$t_{ij} = \frac{1}{c} \int_{s_1}^{s_2} \left( \frac{1}{\mu_i} - \frac{1}{\mu_j} \right) ds, \dots\dots\dots (6)$$

where  $s$  denotes the path (assumed rectilinear and therefore common to both frequencies) between the disturbance (at  $s=s_1$ ) and the observer (at  $s=s_2$ ), and  $\mu_i$  and  $\mu_j$  denote values of  $\mu$  at the two frequencies. Hence if the electron density along the appropriate path is known, the value of  $t_{ij}$  can be calculated from (5) and (6).

We first consider the particular case in which the electron density distribution in the appropriate part of the solar atmosphere (i.e. the corona) is spherically symmetrical and given by Baumbach's empirical formula as revised by Allen(9). We consider only the rays that escape from the direction of the centre of the sun's disk; in this case the path of the rays is radial. Under these conditions, the time intervals  $t_{ij}$  corresponding to different values of  $s_1$  (i.e. different heights of the initiating disturbance) can be derived from curves published by Jaeger and Westfold(7). Corresponding values of  $t_{12}$  and  $t_{23}$  are found to lie on the curve  $OA$  in Figure 6. (The point  $O$  corresponds to  $s_1 = \infty$ ,\* and the point  $A$  to  $s_1$  equal to the lowest coronal height at which radiation of frequency  $f_1$  (75 Mc/s.) can escape.) The lack of agreement between this curve and the experimental points demonstrates that the frequency drift of type III bursts cannot be explained by this mechanism under the assumed conditions. The argument is not significantly modified if non-central rays are taken into account, and the same conclusion is reached.

It is possible, however, that marked local departures from the Baumbach-Allen distribution may occur at times. Generally speaking, the more gradual the electron-density gradient, the greater the time delays between different frequencies. If suitable electron-density distributions are postulated, there is no limit to the value which individual time delays ( $t_{12}$  or  $t_{23}$ ) caused by group retardation may assume. However, it is proved in Appendix I (a) that *there*

\* The observer (at  $s_2$ ) is also regarded as being at infinity for this calculation.

is a fundamental restriction to the ratio  $t_{23}/t_{12}$ , whatever the electron density distribution of the atmosphere may be. It is shown that the ratio cannot exceed the value

$$\left(\frac{t_{23}}{t_{12}}\right)_{max} = \frac{f_2^{-2} - f_3^{-2}}{f_1^{-2} - f_2^{-2}}.$$

In accordance with equation (3), the value of  $(t_{23}/t_{12})_{max}$  appropriate to Figure 6 corresponds to the line marked " $n=3$ " in that diagram. It follows from this theorem that if the frequency drift of a burst were due solely to selective group retardation, the observed value of  $t_{23}/t_{12}$  would lie in the shaded sector of the diagram. This is clearly not the case for the observed type III bursts. It is therefore concluded that *the frequency drift of these bursts is not due solely to selective group retardation.*

It may be objected that the theorem lacks generality because only the case of rectilinear propagation has been considered. When the paths are curved, different frequencies follow different paths, and the problem can only be treated in particular cases. In the case of a spherically symmetrical atmosphere, however, the theorem is shown in Appendix I (b) to be valid for all paths.

(ii) *The Interpretation of Frequency Drift in Terms of the Outward Motion of a Localized Agency in the Solar Atmosphere.*—This mechanism depends on two ideas. The first is that emission of coherent waves in an ionized atmosphere is mainly confined to a narrow frequency range near the frequency  $f_0$  at which the refractive index is zero (10, 7)—such emission corresponds to the peak at  $f_0$  in Figure 5 (a). The second is that the value of  $f_0$  in the solar atmosphere decreases continuously with height—this follows from equation (5) if  $N$  is assumed to decrease continuously with height. Thus according to these ideas, a burst produced by an outward moving agency would show frequency drift, in the characteristic direction, determined by the velocity of the agency and by the electron density distribution of the atmosphere.

We first consider the case of an agency travelling radially outwards through the solar atmosphere towards the observer at a constant velocity  $V$ . Then, assuming the electron density distribution to be given by the Baumbach-Allen formula previously cited, the resulting pairs of time delays,  $t_{12}$  and  $t_{23}$  (see Fig. 6) for different values of  $V$  are found to correspond to points on the line  $OB$ . (The numbers on the line indicate values of  $V$  in units of  $10^4$  km./sec.) Under the conditions assumed, points lying above this line would correspond to acceleration of the agency, and those below to retardation. On this basis, therefore, the majority of the observed bursts would correspond to agencies with slight outward acceleration. It should, however, be stressed that the gradient of the line  $OB$  depends on the assumed electron density distribution.

It follows from Snell's law that radiation originating at a surface of zero refractive index can only escape in the direction normal to that surface. In the case of a spherically symmetrical atmosphere, therefore, the path of the agency considered above (i.e. the radial path directed towards the observer) is the only one from which radiation can escape from the zero-refractive-index level towards the observer. This suggests that, if the present hypothesis is



correct, type III bursts that are observable at the earth would tend to originate at the centre of the sun's disk.

Some support is given to the moving-agency hypothesis by the evidence, given in Section III (c), that the exponential decay constant of type III bursts tends to decrease with decreasing frequency. For it has been suggested by Westfold(11) that the decay constant is to be identified with the frequency of electron-atom collisions in the solar atmosphere at the region of the initiating disturbance. Now on the usual assumptions regarding the electron density and temperature in the solar corona, the collision frequency decreases continuously outwards (see, for instance, Smerd 12). Hence a tendency for the decay constant to decrease with decreasing frequency would imply outward motion of the initiating disturbance.

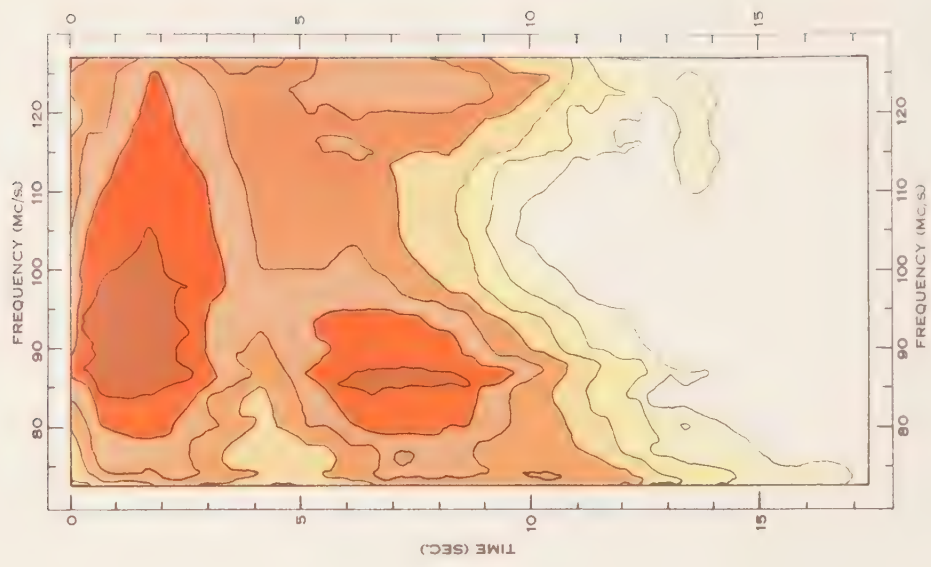
These arguments indicate that the explanation of the frequency drift of type III bursts in terms of the motion of exciting agencies is consistent with the observational evidence. With the assumed electron density distribution, velocities of between about  $2 \times 10^4$  and  $10^5$  km./sec. would be required to account for the observed drift rates (see Fig. 6). Corpuscular streams with these velocities have not been observed in the solar atmosphere, but in any case such streams would be likely to be highly ionized and may consequently escape optical detection. It is noteworthy, however, that particles attain such velocities in terrestrial discharge phenomena, such as lightning flashes. Alternatively, if the burst radiation originated in regions where levels of zero refractive index at different frequencies are abnormally crowded (e.g. at the borders of prominence material), the observed drift could be explained by correspondingly lower velocities of the exciting agency.

A further possibility, pointed out to the author by Dr. D. F. Martyn and Dr. J. C. Jaeger independently, is that a type III burst may be caused by the rapid dispersion of a cloud of gas, situated in the corona, which is initially of considerably higher electron density than the atmosphere surrounding it. Assuming, as before, that most of the emission at any frequency originates at the level of zero refractive index, the decay of electron density in the cloud would cause the frequency of maximum intensity of the burst to decrease with time.

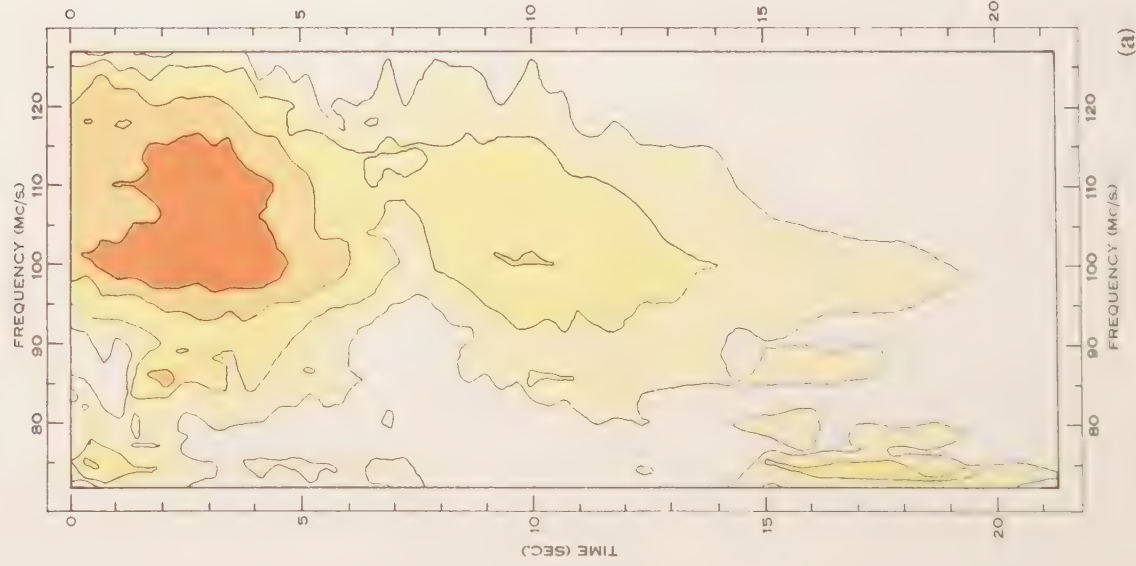
## V. CONCLUSION

These observations represent a first attempt at investigating the spectra of isolated bursts from the sun. They establish that a considerable proportion of the bursts exhibit a distinct type of spectrum (designated in this paper as "type III"). Other bursts appeared to show widely different characteristics, but, in view of the limited range of frequencies so far examined, it seems premature to conclude that several distinct types exist. Observation over a wider frequency range may reveal further systematic properties and provide a more coherent picture.

Although we have a fair knowledge of the properties of these bursts, the physical mechanism responsible for their production is not known. All that can be said at present is that the short lifetime of the bursts suggests an origin in localized sources in the solar atmosphere.



(b)



(a)

correct, type  
at the centre

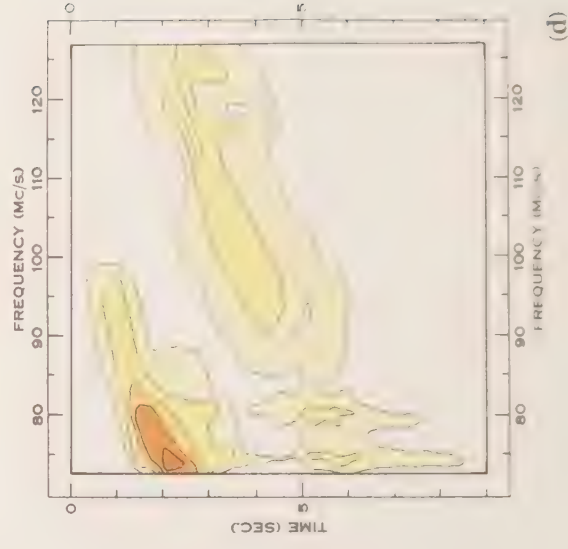
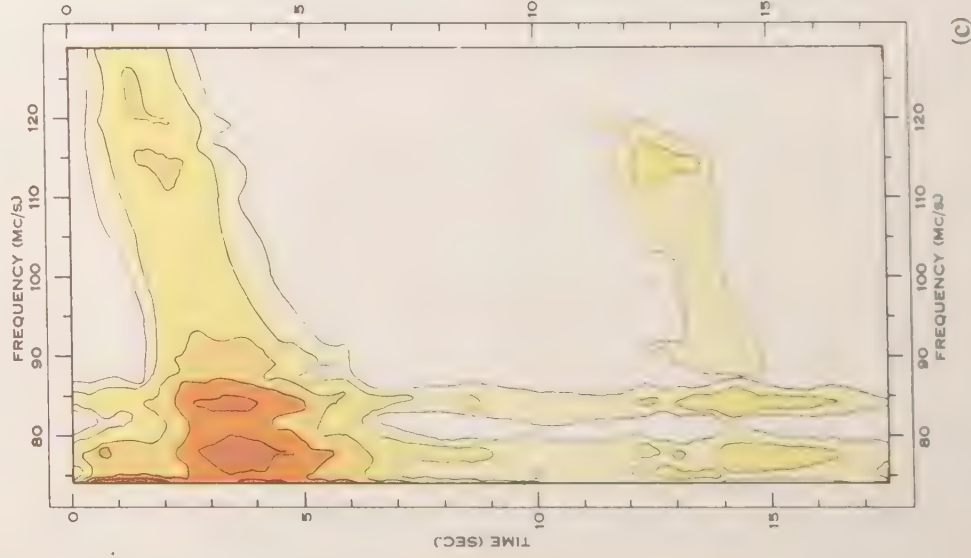
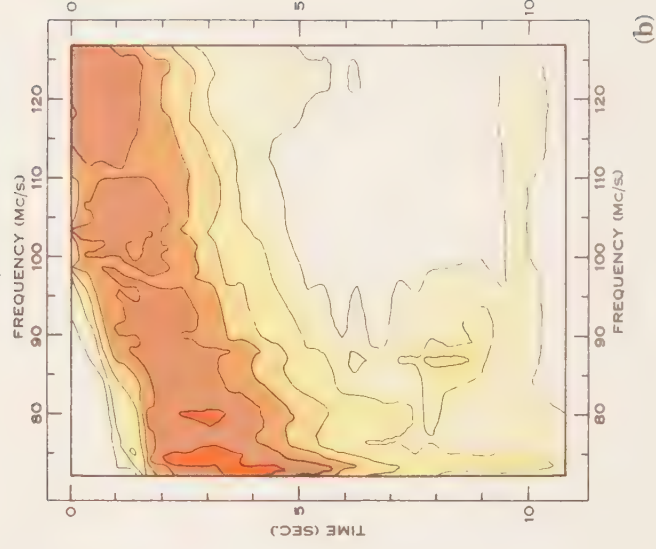
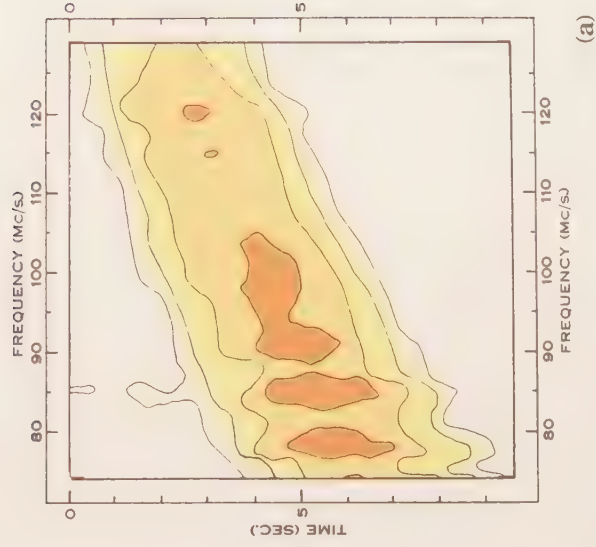
Some su  
given in Sect  
tends to dec  
Westfold(11)  
electron-atom  
disturbance.  
temperature  
outwards (se  
to decrease  
initiating di

These a  
type III bur  
observational  
velocities of  
for the obser  
have not be  
would be li  
detection.  
terrestrial d  
the burst ra  
different fre  
material), t  
velocities of

A furth  
Dr. J. C. Ja  
rapid disper  
considerably  
Assuming,  $\epsilon$   
the level of  
would cause  
time.

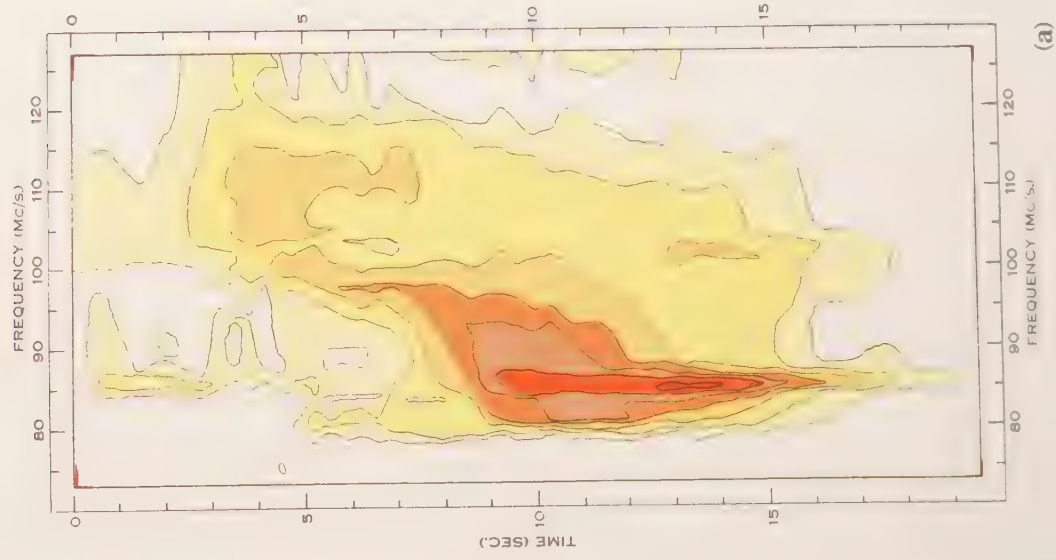
These  
of isolated  
of the burs  
" type III"  
but, in vie  
premature  
wider frequ  
a more col

Althou  
physical m  
can be said  
in localized

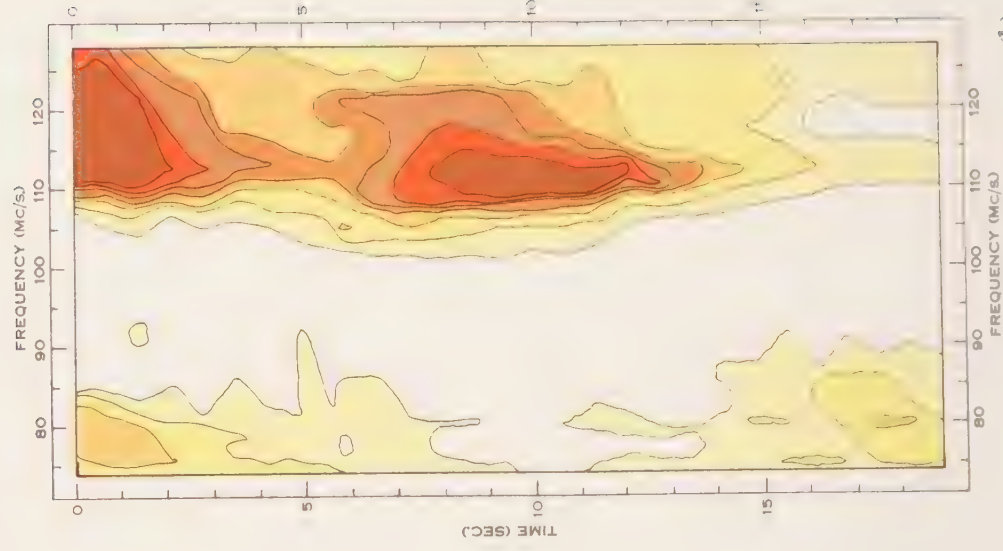




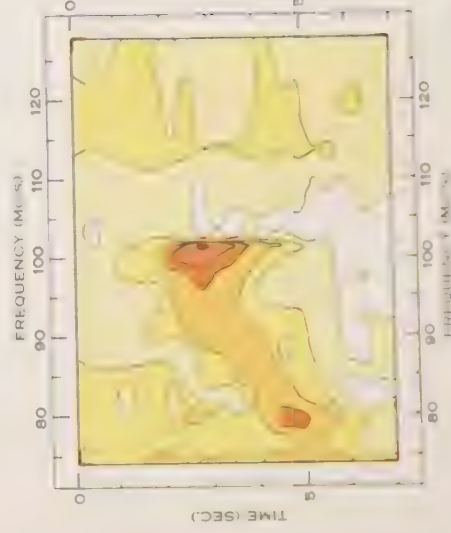




(a)



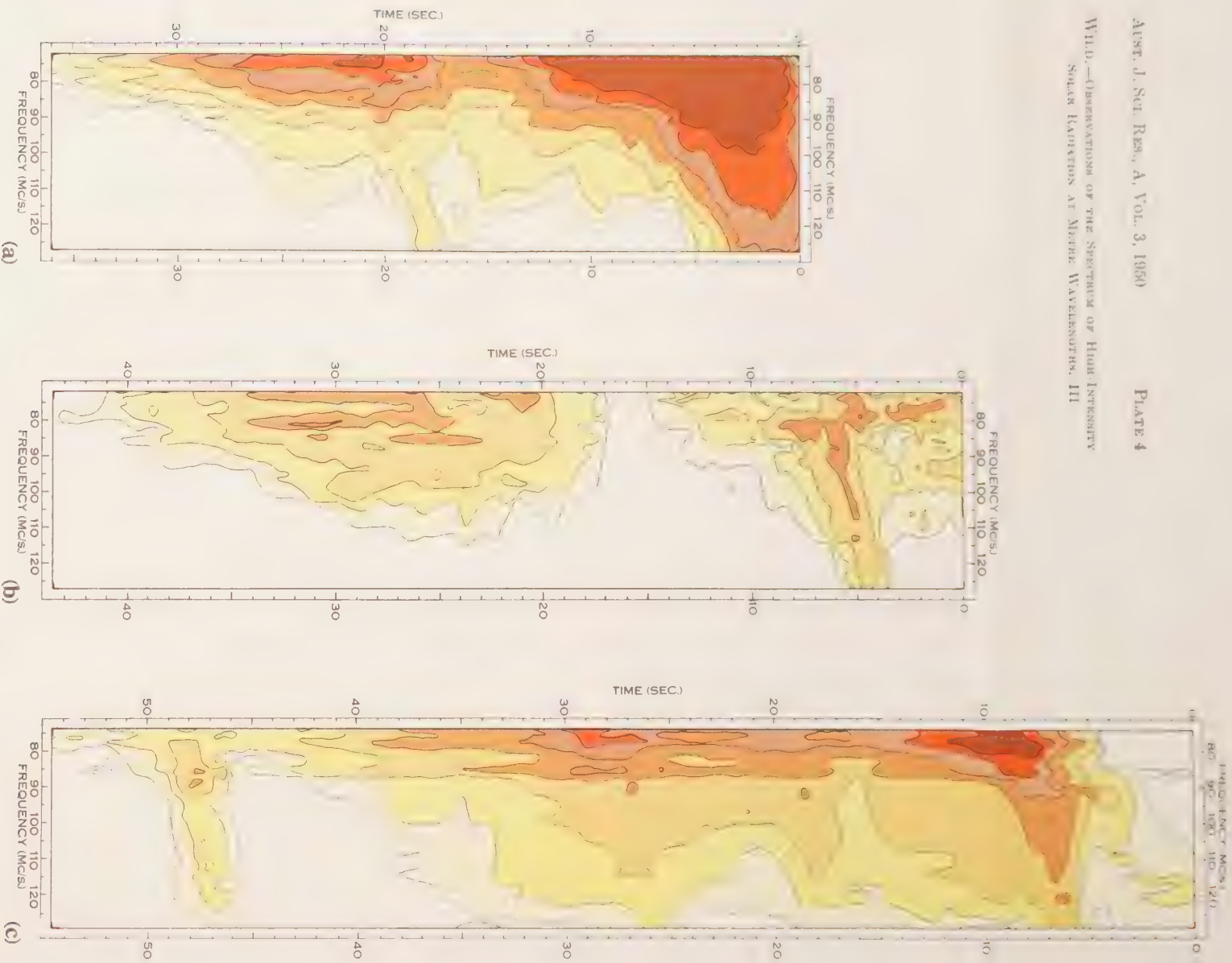
(b)



(c)



WILD.—OBSERVATIONS OF THE SPECTRUM OF HIGH-INTENSITY  
SOLAR EMISSION AT METRE WAVELENGTHS. III







Analysis of the observations shows that the characteristic frequency drift of the type III bursts is not due solely to the selective group retardation experienced by radiation passing through the ionized solar atmosphere, whatever the distribution of ionization might be. The frequency drift may, however, be explained in terms of either the rapid motion, outwards through the solar atmosphere, of a disturbing source, or the rapid dispersion of a relatively dense cloud of gas situated high in the corona.

#### VI. ACKNOWLEDGMENTS

The author is indebted to Mr. L. L. McCready, Mr. J. D. Murray, and Mr. W. C. Rowe, who took part in the observational phase of this work. The author wishes to thank Dr. J. L. Pawsey, Dr. J. C. Jaeger, and Mr. S. F. Smerd for helpful criticism and advice in the preparation of this paper.

#### VII. REFERENCES

- (1) WILD, J. P., and MCCREADY, L. L.—*Aust. J. Sci. Res. A* **3**: 387 (1950).
- (2) WILD, J. P.—*Aust. J. Sci. Res. A* **3**: 399 (1950).
- (3) PAWSEY, J. L.—*Proc. Instn. Elect. Engrs.* **97** (1950).
- (4) PAYNE-SCOTT, RUBY.—*Aust. J. Sci. Res. A* **2**: 214-27 (1949).
- (5) WILLIAMS, S. E.—*Nature* **162**: 108 (1948).
- (6) PAYNE-SCOTT, RUBY, YABSLEY, D. E., and BOLTON, J. G.—*Nature* **160**: 256-7 (1947).
- (7) JAEGER, J. C., and WESTFOLD, K. C.—*Aust. J. Sci. Res. A* **2**: 322-34 (1949).
- (8) JAEGER, J. C., and WESTFOLD, K. C.—*Aust. J. Sci. Res. A* **3**: 376 (1950).
- (9) ALLEN, C. W.—*Mon. Not. R. Astr. Soc.* **107**: 426-32 (1947).
- (10) MARTYN, D. F.—*Nature* **159**: 26-7 (1947).
- (11) WESTFOLD, K. C.—*Aust. J. Sci. Res. A* **2**: 169-83 (1949).
- (12) SMERD, S. F.—A radio-frequency representation of the solar atmosphere. *Proc. Instn. Elect. Engrs.* (in press).

#### EXPLANATION OF PLATES 1-4

Time-frequency-intensity diagrams showing spectra of individual isolated bursts with time variations.

*Note.*—Some of the diagrams (e.g. Plate 2a and 2c) show narrow vertical "ridges" down two or more lines of constant frequency in the low-frequency half of the observed range. This effect, which is never more than one contour in magnitude, is believed instrumental in origin—see Part I.

##### PLATE 1

- (a) 6.iv.49; 01 hr. 27 min. (G.M.T.)
- (b) 21.iii.49; 02 hr. 43 min.

##### PLATE 2

###### Type III bursts

- (a) 31.v.49; 01 hr. 09 min.
- (b) 21.iii.49; 02 hr. 44 min.
- (c) 31.v.49; 00 hr. 38 min.
- (d) 18.iii.49; 05 hr. 13 min.

##### PLATE 3

- (a) 13.iv.49; 00 hr. 40 min.
- (b) 6.v.49; 03 hr. 11 min.
- (c) 2.v.49; 02 hr. 10 min.

##### PLATE 4

- (a) 11.iv.49; 05 hr. 30 min.
- (b) 11.iv.49; 05 hr. 59 min.
- (c) 30.v.49; 01 hr. 13 min.

## APPENDIX I

(a) The following theorem is proved\* (see Section IV (b) (i)) :

If  $t_{ij}$  denotes the difference in the time taken by waves of frequency  $f_i$  and  $f_j$  to travel along a common path between two points in an ionized medium (in general unhomogeneous), and if the refractive index of the medium is that given by Lorentz's theory, then the ratio  $t_{23}/t_{12}$  for three frequencies ( $f_1 < f_2 < f_3$ ) satisfies the condition

$$\frac{t_{23}}{t_{12}} \leq \frac{f_2^{-2} - f_3^{-2}}{f_1^{-2} - f_2^{-2}}.$$

The time taken by a wave of frequency  $f$  to travel between points  $s_1$  and  $s_2$  is

$$t = \frac{1}{c} \int_{s_1}^{s_2} \frac{1}{\mu} ds, \dots\dots\dots (\text{I.1})$$

where the refractive index,  $\mu$ , is given by equation (5).

The electron density,  $N$ , (equation (5)) is a function of path distance,  $s$ , so for a particular path we can write

$$\mu^2 = 1 - \frac{\varphi(s)}{f^2}, \dots\dots\dots (\text{I.2})$$

where  $\varphi(s)$  is a positive continuous function independent of  $f$ . We are concerned only with those paths along which waves of frequency  $f$  can be propagated, and demand that  $1 \geq \mu \geq \varepsilon$  in  $(s_1, s_2)$ , where  $\varepsilon$  is a positive number which may be chosen as small as desired, i.e.

$$0 \leq \frac{\varphi(s)}{f^2} \leq 1 - \varepsilon^2 \text{ in } (s_1, s_2). \dots\dots\dots (\text{I.3})$$

Writing  $z = \frac{1}{f^2}$ , equations (I.1), (I.2), and (I.3) give

$$t = \frac{1}{c} \int_{s_1}^{s_2} \{1 - z\varphi(s)\}^{-\frac{1}{2}} ds, \quad 0 \leq z\varphi(s) \leq 1 - \varepsilon^2 \text{ in } (s_1, s_2).$$

Therefore

$$\frac{dt}{dz} = \frac{1}{2c} \int_{s_1}^{s_2} \{1 - z\varphi(s)\}^{-3/2} \varphi(s) ds \geq 0,$$

and

$$\frac{d^2t}{dz^2} = \frac{3}{4c} \int_{s_1}^{s_2} \{1 - z\varphi(s)\}^{-5/2} \{\varphi(s)\}^2 ds \geq 0.$$

Thus  $t$  increases with  $z$  and the curve of  $t$  against  $z$  has its centre of curvature above the curve (see Fig. 7). Then, if  $f_1, f_2, f_3$  are three frequencies such that  $f_1 < f_2 < f_3$ , and if  $z_1, z_2, z_3$  and  $t_1, t_2, t_3$  are the corresponding values of  $z$  and  $t$ , it follows from the geometry of the figure that

$$\frac{t_2 - t_3}{t_1 - t_2} \leq \frac{z_2 - z_3}{z_1 - z_2}.$$

i.e.

$$\frac{t_{23}}{t_{12}} = \frac{t_2 - t_3}{t_1 - t_2} \leq \frac{f_2^{-2} - f_3^{-2}}{f_1^{-2} - f_2^{-2}}. \dots\dots\dots (\text{I.4})$$

\* The proof given here is due to Dr. J. C. Jaeger (personal communication), being considerably simpler and more elegant than one originally drafted by the author.

(b) We now show that if  $t_{ij}$  denotes the difference in time taken by waves of frequency  $f_i$  and  $f_j$  to travel along paths (in general different) from a point in a spherically symmetrical ionized atmosphere to a distant observer outside that atmosphere, then condition (I.4) is satisfied for all possible paths.

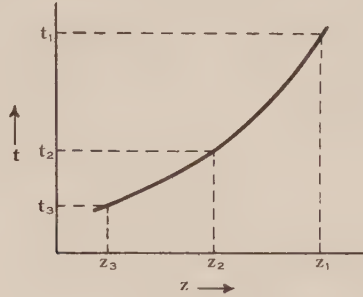


Fig. 7

For a discussion on the trajectories of rays in a spherically symmetrical atmosphere the reader is referred to Jaeger and Westfold(8). A path that reaches the distant observer will emerge from the atmosphere parallel to (at a distance  $d$ , say) from the line joining the observer and the centre of the atmosphere. The equation to this path is given by

$$ds = \left(1 - \frac{d^2}{r^2\mu^2}\right)^{-\frac{1}{2}} dr,$$

where  $r$  denotes radial distance.

Thus from (I.1) we have

$$t = \frac{1}{c} \int_{r_1}^{r_2} (\mu^2 - d^2/r^2)^{-\frac{1}{2}} dr, \quad \left(\mu > \frac{d}{r}\right).$$

Since the electron density distribution is spherically symmetrical, we may write (cf. (I.2))

$$\mu^2 = 1 - \frac{\varphi(r)}{f^2},$$

so that

$$t = \frac{1}{c} \int_{r_1}^{r_2} \{1 - z\varphi(r) - d^2/r^2\}^{-\frac{1}{2}} dr,$$

where  $z = \frac{1}{f^2}$ .

Again we find that  $\frac{dt}{dz}$  and  $\frac{d^2t}{dz^2}$  are positive (or zero) for paths along which waves can be propagated  $\left(1 \geq \mu \geq \frac{d}{r} + \varepsilon\right)$  in  $(r_1, r_2)$ , and so condition (I.4) follows as before.



# POLAROGRAPHY WITH ALTERNATING CURRENTS

## I. OUTLINE OF THEORY, APPARATUS, AND TECHNIQUE

By B. BREYER,\* F. GUTMAN,† and S. HACOBIAN\*

[Manuscript received July 20, 1950]

### Summary

A method of polarographic analysis employing the superposition of a small sinusoidal alternating field of low frequency upon the direct potential applied to the dropping mercury electrode has been further developed. Equations are given for the calculation of the ionic concentration from the alternating current (maximum differential current) at the characteristic "summit potential" (analogous to  $E_{\frac{1}{2}}$  in conventional polarography). This current is proportional to the concentration of the ion species undergoing discharge. The advantages of the method over conventional polarography can be summarized as follows:

- (i) A.C. polarography can be carried out in solutions containing dissolved oxygen, i.e. in air.
- (ii) Tedious curve plotting is substituted by a single current reading at the summit potential which yields directly both the half-wave potential and the concentration of the electro-reducible substance.
- (iii) Small amounts of less noble ions can be estimated even in the presence of a large excess of nobler ions.
- (iv) Polarographic waves only 40 mV. apart are clearly separable.
- (v) An improved all round reproducibility is obtained.
- (vi) There is a simplified system of recalibration in case of capillary replacement.
- (vii) The use of a delicate galvanometer is eliminated.

## I. INTRODUCTION

In earlier papers published by two of us(1-4) the theoretical and experimental foundations of polarographic analysis employing the superposition of a small sinusoidal alternating voltage upon the direct potential applied to the dropping mercury electrode, were laid. The treatment referred to an electrode which, on increasing the applied direct potential, passes from polarization through a reversible stage into a state of concentration polarization.

Electrode processes involving the application of an alternating field of low frequency have been investigated by various workers(5-9) from time to time, the usual measuring unit being the Cathode-Ray Oscillograph. The method of polarographic analysis described in this paper is considerably simpler than oscillographic polarography and, unlike the latter (see, for example, Delahay 6), yields results of good reproducibility.

The method used is shown in Figure 1.

\* Physico-Chemical Laboratory, Department of Agriculture, University of Sydney.

† Department of Electrical Engineering, New South Wales University of Technology, Sydney.

It consists of the superposition of a small alternating potential  $V$  (1–90 mV.) on a D.C. potential applied by the potentiometer  $P_1$  to the dropping mercury

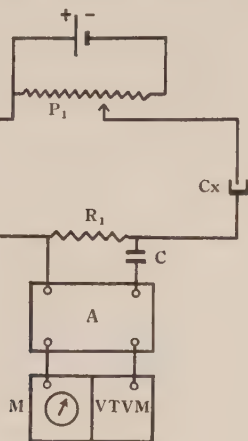


Fig. 1 (a).—Diagram of circuit.

$P_1$ , potentiometer;  $V$ , sinusoidal e.m.f. 1–90 mV.;  $R_1$ , voltage-dropping resistor;  $C$ , condenser;  $C_x$ , polarographic cell;  $A$ , amplifier; VTVM, vacuum tube voltmeter;  $M$ , microammeter.

electrode  $C_x$ . The alternating current flowing at a given D.C. potential is amplified by the amplifier  $A$  and the amplified current read by means of a vacuum tube voltmeter arrangement.

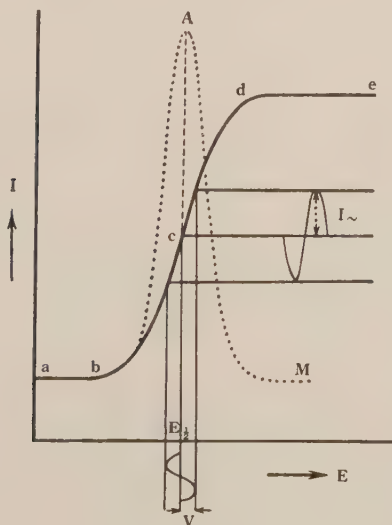


Fig. 1 (b)

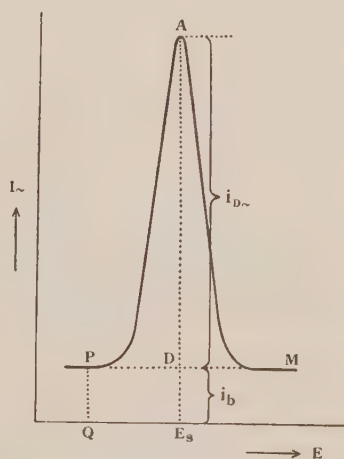


Fig. 1 (c)

The polarogram obtained by plotting the deflection of the meter  $M$  against applied D.C. potential is shown schematically in Figure 1 (c). This type of

curve we shall call a "differential polarographic wave" because it results from a differential operation applied to the reversible D.C. polarographic step as shown in Figure 1 (b). A small alternating voltage of amplitude  $V$  is applied to the non-linear current-voltage portion of the conventional polarogram with the result that a corresponding alternating current is obtained with the same frequency but a varying amplitude  $I \sim$  depending on the applied D.C. voltage.

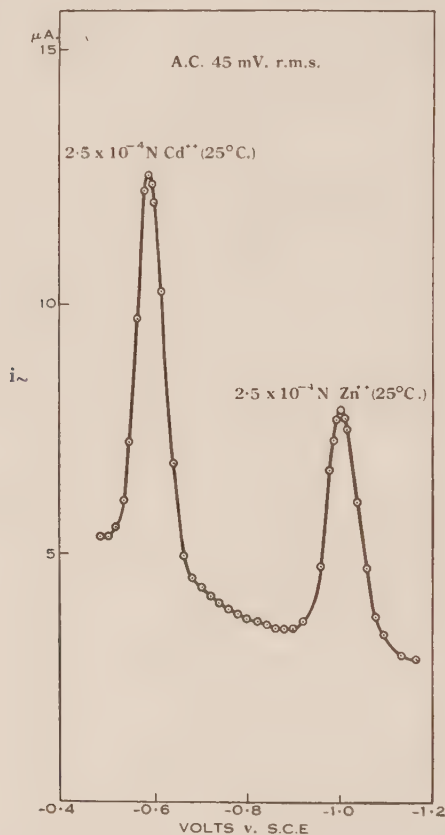


Fig. 2.—Polarogram of  $2.5 \times 10^{-4} N \text{ Cd}^{++}$  and  $2.5 \times 10^{-4} N \text{ Zn}^{++}$  at  $25^{\circ}C$ . Supporting electrolyte:  $0.1M$  sodium acetate +  $0.1M$  acetic acid (pH 4.7) + dissolved air.  $m$ ,  $1.46 \text{ mg./sec.}$   $t$ ,  $3.2 \text{ sec.}$  at zero applied potential. Ohmic resistance of cell circuit,  $1385 \Omega$ . Potentiometer resistance,  $1000 \Omega/V$ .

At the flat portion of the curve corresponding to  $ab$  and  $de$ , where the D.C. current is virtually constant,  $dI/dE$  nearly vanishes, i.e. no differential current is observed. It follows from the diagram that a maximum of alternating current will occur at the half-wave potential  $E_{\frac{1}{2}}$  (dotted curve  $bAM$ , Fig. 1 (b)). We shall call  $A$  the "summit potential" (abbreviated  $E_s$ ) of the differential wave  $PAM$  (Fig. 1 (c)). This corresponds to the ordinary half-step potential.  $AD$  represents the magnitude of the differential current at the summit potential and can be represented symbolically by  $i_{D\sim}$ .  $PQ$  or  $DE_s$  represents the "condenser" or "base-current" which depends upon the nature of the supporting

electrolyte and the capillary used.  $AD$  is proportional to the ion concentration. A typical polarogram of cadmium and zinc is shown in Figure 2. The maximum differential current occurs at the half-wave potential as was shown theoretically (1, 4).

The differential current ( $i_{D\sim}$ ) during depolarization is represented by the equation

$$i_{D\sim} = V \cos \omega t (j\omega C + 1/R_D), \quad \dots \dots \dots (1)$$

where  $j = D/\omega$  and  $D \equiv d/dt$ ,

$V$  = peak alternating potential,

$\omega$  = circular frequency,

and  $C = C_H + C_D$ ,

where  $C_H$  is the capacity of the Helmholtz-Gouy double layer and  $C_D$  is the dynamic capacitance corresponding to the dynamic resistance  $R_D$  (for details see 1, 3, 4).

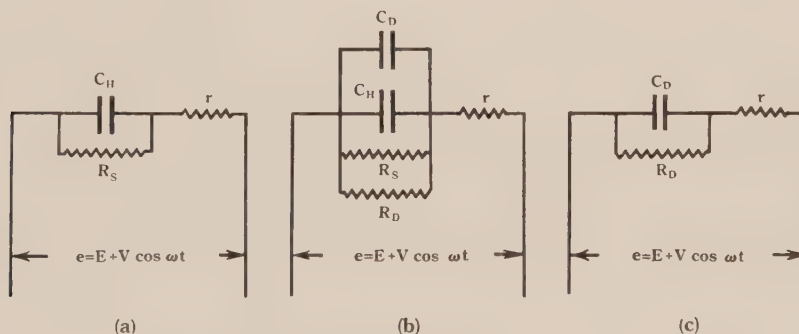


Fig. 3

Equation (1) does not take into account the ohmic resistance of the cell which is assumed to be zero. In actual practice the ohmic resistance is never zero and the equivalent circuit of the depolarized electrode can be represented by Figure 3 (b).

It has been shown (1, 3, 4) that at the half-wave potential  $C_D$  reaches a maximum and  $R_D$  a minimum. Since at this point  $R_D$  ordinarily is  $\ll R_s$  and  $C_D$  ordinarily is  $\gg C_H$ , the conditions operating can be represented by the diagram in Figure 3 (c). Figure 3 (a) represents diagrammatically the polarized electrode. The magnitude of  $C_H$  at the summit potential is given by

$$C_H = i_b / \omega (V_{r.m.s.}^2 - r^2 i_b^2)^{\frac{1}{2}}, \quad \dots \dots \dots (2)$$

where  $V_{r.m.s.}$  = r.m.s. voltage,

$i_b$  = base current corresponding to the summit potential (see

Fig. 1 (c)),

and  $r$  = ohmic resistance of the cell circuit.

The differential quantities  $C_D$  and  $R_D$  are expressed by equations (3) and (4) respectively. Both  $C_D$  and  $R_D$  are directly proportional to the concentration.

$$C_D = C_o n^2 F^2 (mt)^{\frac{1}{2}} d / 2.35 RT, \quad \dots \dots \dots (3)$$



where  $C_{ox}$  = concentration of oxidant in moles/litre,

$n$  = number of electrons involved in the reduction,

$F$  = Faraday,

$m$  = weight of mercury flowing per second from the mercury dropping electrode in grams per second,

$t$  = drop time in seconds,

$R$  = gas constant,

$T$  = absolute temperature,

and  $d$  = thickness (in cm.) of ionic space charge within which the electrochemical process takes place.

At the summit potential the dynamic resistance is given by

$$R_D = \frac{4RT}{nF i_d}, \quad \dots \dots \dots (4)$$

where  $i_d$  is the diffusion current.

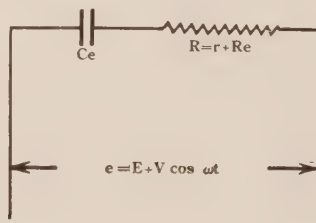


Fig. 4

The parallel circuit  $C_D$ - $R_D$  given in Figure 3 (c) can be transformed into the equivalent series arrangement as shown in Figure 4, by means of the well-known relationships

$$C_e = C(1 + 1/\omega^2 C^2 R_D^2), \quad \dots \dots \dots (5)$$

$$R_e = R_D/(1 + \omega^2 C^2 R_D^2), \quad \dots \dots \dots (6)$$

and

$$R' = R_e + r. \quad \dots \dots \dots (7)$$

These equations make it possible to calculate the differential current from a knowledge of  $C$ ,  $R_D$ , and  $r$ , where  $C$  equals the sum of  $C_H$  and  $C_D$  as shown before.

Thus the impedance,  $|Z|$ , of the whole network can be calculated from

$$|Z| = (R'^2 + |X|_{ce}^2)^{\frac{1}{2}}, \quad \dots \dots \dots (8)$$

where  $|X|_{ce} = 1/\omega C_e$ ,

and the total alternating current passing through the cell at the summit potential can be represented by

$$i_{D\sim} = \frac{V_{r.m.s.}}{|Z|} - i_b. \quad \dots \dots \dots (9)$$

Since  $i_{D\sim}$  can be found experimentally and  $C_H$  and  $R_D$  can be calculated from a knowledge of  $r$ ,  $i_d$ , and  $i_b$  respectively,  $C$  can be evaluated which in turn gives the value of  $C_D$  without the use of expression (3).

Neglecting the ohmic resistance  $r$  and  $C_H$ , the impedance of the cell becomes

$$|Z| = (R^2 + 1/\omega^2 R_D^2)^{\frac{1}{2}},$$

$$= (a/C_{ox}^2 + b/C_{ox}^2)^{\frac{1}{2}} = \frac{K}{C_{ox}}, \quad \dots\dots\dots (10a)$$

where  $a$ ,  $b$ , and  $K$  are constants independent of the concentration of the oxidant. Equation (9) then becomes

$$i_D \sim \frac{V_{r.m.s.} C_{ox}}{K} - i_b, \quad \dots\dots\dots (10b)$$

which shows that the maximum differential current at the summit potential is linearly proportional to the concentration of the reducible ion undergoing the electro-reduction process. Both  $C_H$  and  $r$ , which do not enter equations (10a) and (10b), produce a departure from the linearity of the above relationship. Provided the ohmic resistance,  $r$ , of the cell circuit is kept low, the departures from linearity are small and do not in any way prejudice the accuracy of the analytical results obtained.

The quantity  $d$  appearing in equation (3) depends on the nature of the supporting electrolyte present and varies with the nature of the ion species undergoing discharge.

It should be stressed that the theoretical treatment applies only to reversible electrode processes. In irreversible processes the D.C. polarographic step deviates more or less from the characteristic  $S$ -shaped form, the rising part approaching more and more a straight line. Here  $dI/dE$  becomes constant and hence no A.C. wave appears.

## II. CIRCUITS AND APPARATUS

The D.C. potential applied to the cell  $C_x$  is obtained from the accurate potentiometer  $P_1$  (resistance 40 ohms) as shown in Figure 5. Superposed upon this direct voltage is a small alternating potential derived from the potential divider  $P_2$ . An A.C. signal of power frequency (50 c/s.) taken from an external 2.5 volt secondary transformer  $T_1$  was employed via a magnetic constant voltage regulator  $M$ . Setting of  $P_2$  permits the choice of different values of alternating voltage between 1 and 90 mV. The alternating current is measured by the voltage drop across the adjustable resistor  $R_1$ . Twelve values of  $R_1$  can be selected ranging from 10 to 500 ohms. Selection of  $R_1$  determines the sensitivity of the instrument, according to the concentration range desired. The voltage across  $R_1$  is applied to a conventional audio-frequency amplifier (gain approximately 10,000) through a condenser  $C_1$  which filters out the D.C. component.

The ease of electronic amplification permits the introduction of a peak smoothing device in the form of a peak reading valve voltmeter of proper time constant which completely smoothes out the unwanted current oscillations due to the dropping electrode.

The alternating output of the amplifier appears across the load resistor  $R_2$  and charges the large condenser  $C_3$  through the diode rectifier  $V_1$  to a potential corresponding very nearly to the peak value of the instantaneous current through the polarographic cell. When the current subsequently decreases, due to a

detachment of a mercury drop, the diode becomes non-conducting and prevents the discharge of  $C_3$  through the amplifier output circuit.  $C_3$ , however, is allowed to discharge through the resistor  $R_3$ . If the time constant  $R_3C_3$  of the combination is of the same order as the drop time of the capillary, all unwanted oscillations disappear. By varying the magnitude of  $R_3$  the time constant can be regulated at will. The voltage across  $C_3$  is applied to the grid of a conventional D.C. single stage valve voltmeter employing two 6F6 valves in a balanced degenerative circuit. The zero is set by means of the potentiometers  $P_3$  and  $P_4$  which control the grid and anode potentials respectively. Coarse and fine controls are provided. The employment of the peak smoothing circuit permits the operation

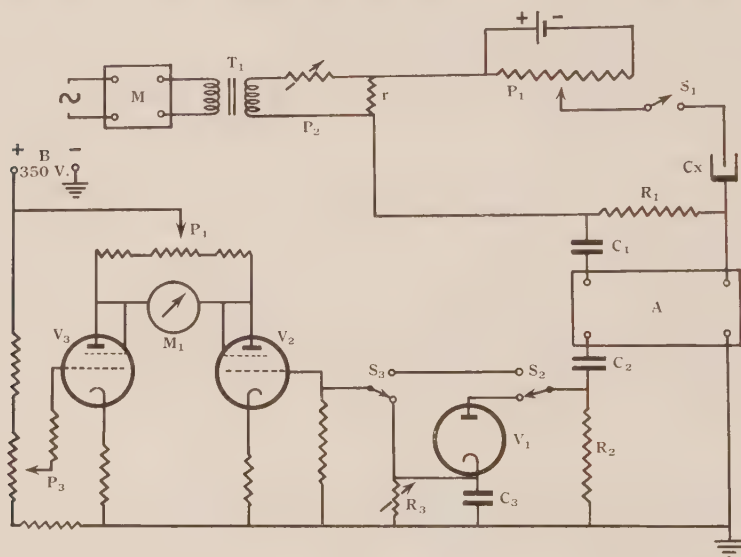


Fig. 5

of the instrument with a "suppressed zero", i.e. the use of a high gain with subsequent balancing out of the residual capacity or base current by adjusting the meter zero. This current flows through the polarographic cell, even in the absence of a dischargeable substance. The ganged switches,  $S_2$ ,  $S_3$ , permit the peak smoothing arrangement to be disconnected from the circuit if so desired.

The alternating signal voltage and all supply voltages are obtained from three separate external transformers and two rectifiers connected to the mains via a magnetic voltage stabilizer. Opening of the switch  $S_1$  permits the checking of the electrical zero of the instrument.

Changes in the capillary characteristics can easily be allowed for by varying the gain, that is, by calibrating the instrument against a standard solution and adjusting the amplification till a standard reading is obtained. This is a considerable convenience in practical work.

No highly sensitive and necessarily delicate galvanometer is needed, since it is easy to obtain the readings on a moving coil meter of, say, 100  $\mu$ A. full-scale

sensitivity. Any commercial recorder can be connected to the output to operate a pen recorder and/or to allow remote indication.

### III. EXPERIMENTAL TECHNIQUE

Whereas in ordinary polarography it is necessary to plot the current-voltage curve of the discharge, and subsequently to evaluate the graph, in the present method it is sufficient merely to adjust the direct current potentiometer  $P_1$  until  $M_1$  (Fig. 5) shows a maximum reading, which takes place at the half-wave potential (summit potential). This maximum differential current reading when referred to a previously prepared calibrating chart, yields directly the concentration of the substance. The half-wave potential is equal to the summit potential at this point.

The method of analysis is as follows: The instrument is first "standardized" against a reference solution of known composition and concentration employing the same supporting electrolyte as the one to be used in actual analysis. This standardization, whilst not absolutely necessary, leads to a reproducibility of results substantially better than obtained with ordinary polarography. Thus the reproducibility when employing  $10^{-3}N$  cadmium solution in either  $0.1N$  KCl or  $0.1N$  HCl as supporting electrolyte was of the order of  $\pm 1$  per cent. The cell should be thermostated to within  $\pm 1^\circ C$ ., the temperature coefficient of the A.C. wave height being about  $0.4$  per cent./ $^\circ C$ . (see Part II of this series).

The alternating current method is considerably faster than conventional polarography, the time needed for the standardization being not more than that necessary for the calibration of, e.g. a glass electrode. In order to carry out the standardization, the test solution of known concentration is introduced into the electrolysis cell and the gain of the amplifier adjusted in such a way that the meter  $M_1$  indicates a fiducial reading at the summit potential. The standardizing solution is now replaced by the unknown solution, and the meter  $M_1$  read again.

With the use of a low impedance mercury pool anode and  $15$  mV. r.m.s. superposed A.C. voltage, the magnitude of the alternating differential current at the half-wave or summit potential of a particular ion species is substantially larger than the corresponding D.C. diffusion current (see Part II, Table 5). This is a decisive factor in the reduction of error in current readings when A.C. and conventional polarography are compared.

Since the A.C. polarogram of a particular ion species is unaffected by the preceding discharge of other ions, small amounts of less noble metals can be estimated in the presence of an excess of nobler ones (e.g.  $Zn^{++}$  in a solution containing an excess of  $Cd^{++}$  or small amounts of  $Cd^{++}$  in presence of an excess of  $Pb^{++}$  etc.). Thus, for instance, mixtures containing cadmium and zinc in the ratio  $100 : 1$  have been successfully analysed by the new polarographic technique (for details see Part II of this series). In conventional polarography, however, the relatively small step produced by the discharge of the less noble ion is completely masked due to the large preceding step. Shunting of the galvanometer(10) has been proposed to alleviate this but has proved to be only a partial



remedy(11). Here A.C. polarography is particularly valuable because analytical mixtures of the composition outlined above cannot be investigated successfully with the ordinary polarographic method.

Two polarographic waves whose half-wave potentials are separated by only 40 mV. (e.g. those of indium and cadmium in 0.1N HCl) can be distinguished from each other by A.C. polarography as will be shown in Part II of this series. In D.C. polarography a difference of at least 60 mV. between two half-wave potentials is necessary in order that two waves may be just seen and at least 90 mV. separation is needed in order to evaluate the individual step heights(12).

Since oxygen is irreversibly reduced at the dropping electrode, no A.C. oxygen wave is obtained so that A.C. polarography can be carried out in the presence of air. This means a considerable saving of time.

Amplifier calibration should be carried out periodically and linearity between current and meter deflection should always be observed. Since the signal voltage frequency is of the order of the power frequency, proper screening of apparatus and leads is essential to obtain good reproducible results.

The most convenient concentration range for analytical purposes in A.C. polarography is between  $10^{-3}$  and  $10^{-4}$ N. The present limit of sensitivity is approximately  $10^{-5}$ N.

#### IV. ACKNOWLEDGMENT

The authors wish to express grateful acknowledgment to the Trustees of the Science and Industry Endowment Fund for financial assistance.

#### V. REFERENCES

- (1) BREYER, B., and GUTMAN, F.—*Trans. Faraday Soc.* **42** : 645 (1946).
- (2) BREYER, B., and GUTMAN, F.—*Trans. Faraday Soc.* **42** : 650 (1946).
- (3) BREYER, B., and GUTMAN, F.—*Trans. Faraday Soc.* **43** : 785 (1946).
- (4) BREYER, B., and GUTMAN, F.—*Faraday Soc. Disc. No. 1* : 19 (1947).
- (5) BOEKE, J., and VAN SUCHTELEN, H.—*Philips Tech. Rev.* **4** : 231 (1939).
- (6) DELAHAY, P.—*J. Phys. and Coll. Chem.* **53** : 1279 (1949).
- (7) HEYROVSKY, J.—*Faraday Soc. Disc. No. 1* : 212 (1947).
- (8) MATHESON, L. A., and NICHOLS, N.—*Trans. Amer. Electrochem. Soc.* **73** : 193 (1948).
- (9) RANDLES, J. E. B.—*Analyst* **72** : 301 (1947).
- (10) LINGANE, J. J., and KERLINGER, H.—*Industr. Engng. Chem. (Anal. Ed.)* **12** : 750 (1940).
- (11) KOLTHOFF, I. M., and LINGANE, J. J.—“Polarography.” 1st Ed. p. 233. (Interscience Publishers Inc. : New York, 1941.)
- (12) MÜLLER, O. H.—*J. Chem. Educ.* **18** : 320 (1941).

# POLAROGRAPHY WITH ALTERNATING CURRENTS

## II. A.C. POLAROGRAPHY OF CADMIUM, ZINC, LEAD, THALLIUM, INDIUM, BISMUTH, AND ANTIMONY

By B. BREYER,\* F. GUTMAN,† and S. HACOBIAN\*

[Manuscript received July 20, 1950]

### Summary

The A.C. polarographic waves of cadmium, zinc, lead, thallium, indium, bismuth, and antimony are discussed. A number of factors affecting the magnitude of the maximum differential current,  $i_{D\sim}$ , e.g. the effect of series resistance in the cell circuit, amplitude of the A.C. voltage, variation in temperature, variation of supporting electrolyte, presence or absence of air, etc., have been determined.

A solution containing a mixture of bismuth, antimony, lead, thallium, indium, cadmium, and zinc was analysed with A.C. polarography and it was found possible to distinguish each wave of the seven elements simultaneously present. This is not possible in conventional polarography.

### I. INTRODUCTION

In Part I of this series it was shown that alternating current maxima occur at the summit potential ( $E_s$ ) when an ion species is discharged reversibly at the dropping mercury electrode.

This paper deals with the practice of A.C. polarography of seven elements, grouped together because they all can be examined in a common supporting electrolyte.

### II. CADMIUM AND ZINC

In 0.1N potassium chloride (air-free) as supporting electrolyte cadmium and zinc gave well-defined A.C. waves due to the reversible reduction of the hydrated  $Zn^{++}$  and  $Cd^{++}$  ions; the summit potentials corresponded to their half-wave potentials:  $-0.60$  and  $-1.05$  V. (*v. S.C.E.*) for cadmium and zinc ions respectively, in accordance with the reported D.C. half-step potentials(1).

In the presence of air, however, the zinc A.C. wave disappeared, but reappeared on expulsion of dissolved oxygen. The cadmium wave height in air was slightly reduced.

The disappearance of the zinc wave is due to the presence of hydroxyl ions formed by reduction of dissolved oxygen at the dropping electrode(2). The hydroxyl ions formed combine with the  $Zn^{++}$  to produce the irreversibly reduced zincate ion. As irreversible reduction processes do not yield an A.C.

\* Physico-Chemical Laboratory, Department of Agriculture, University of Sydney.

† Department of Electrical Engineering, New South Wales University of Technology, Sydney.

polarogram, the zinc wave disappears altogether. These assumptions are borne out by the following experiments :

- (i) The A.C. zinc wave (in 0.1N potassium chloride) disappears even in the absence of air after the addition of a few drops of concentrated ammonia solution, the A.C. cadmium wave being but little affected under these conditions.
- (ii) Both zinc and cadmium waves were obtained in D.C. polarograms of the same solution.

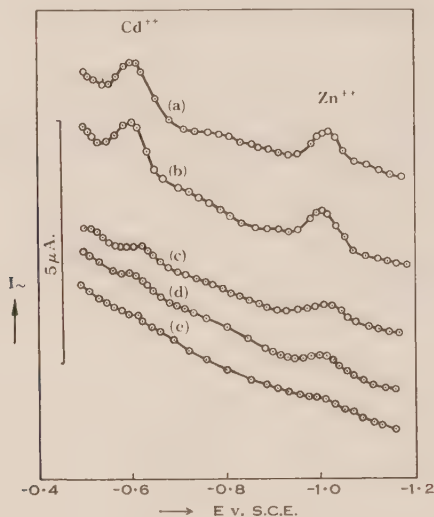


Fig. 1.—A.C. polarograms of  $\text{Cd}^{++}$  and  $\text{Zn}^{++}$ .

- (a)  $2 \times 10^{-5}\text{N}$   $\text{Cd}^{++} + 2 \times 10^{-5}\text{N}$   $\text{Zn}^{++}$  (in air).
- (b) Same as (a) (air-free).
- (c)  $5 \times 10^{-6}\text{N}$   $\text{Cd}^{++} + 5 \times 10^{-6}\text{N}$   $\text{Zn}^{++}$  (in air).
- (d) Same as (c) (air-free).
- (e) Supporting electrolyte, i.e. buffer pH 4.7 (in air).

- (iii) Plotting  $\log (i_d - i)/i$ , i.e.  $\log \frac{[\text{Ox}]}{[\text{Red}]}$ , against  $E$  disclosed that under the above conditions cadmium was reversibly and zinc irreversibly reduced.
- (iv) In  $2 \times 10^{-3}\text{N}$  sodium hydroxide solution a precipitate of cadmium hydroxide was formed and the A.C. waves of both  $10^{-3}\text{N}$   $\text{Cd}^{++}$  and  $10^{-3}\text{N}$   $\text{Zn}^{++}$  disappeared. In D.C. polarography only the cadmium step disappeared whereas zinc gave an irreversible reduction step.
- (v) The oxygen concentration in the solution and, therefore, the concentration of the hydroxyl ions produced cathodically are practically constant. The lowering of the cadmium wave height in presence of oxygen is probably due to the formation of some insoluble cadmium hydroxide at the electrode. It would then follow that the difference between A.C. wave heights obtained in solution containing air, compared

with those obtained in the absence of air, should be the greater the higher the dilution of cadmium. This reasoning is borne out by experiment (cf. Table 1).

- (vi) In 0.1M sodium acetate-acetic acid buffer (pH 4.7) no hydroxo-complex of zinc can be formed. Under these conditions both cadmium and zinc gave well-defined A.C. waves even in the presence of air. The summit potentials for cadmium and zinc ( $10^{-3}\text{N}$ ) were  $-0.60$  and  $-1.005$  V. (v. S.C.E.) respectively in accordance with the  $E_{\frac{1}{2}}$  reported(1).

It can be seen from Figure 1 that A.C. polarography of even dilute cadmium and zinc solutions can be carried out in the presence of air if the pH of the solution is kept suitably low.

TABLE 1  
DIFFERENCE BETWEEN A.C. WAVE HEIGHTS IN PRESENCE AND ABSENCE OF AIR

Solution	$i_{D\sim}$ in Air-Free Solution ( $\mu\text{A.}$ )	$i_{D\sim}$ in Air ( $\mu\text{A.}$ )	Difference (%)
$2 \times 10^{-3}\text{N Cd}^{++}$	13.5	13.6	<1
$1 \times 10^{-3}\text{N Cd}^{++}$	11.3	10.0	11.5
$5 \times 10^{-4}\text{N Cd}^{++}$	9.2	6.2	32.5
$2.5 \times 10^{-4}\text{N Cd}^{++}$	6.3	4.1	35

The determination of small concentrations of zinc ions in presence of a large excess of cadmium ions is a comparatively simple matter in A.C. polarography. Figure 2 shows the differential wave of  $8 \times 10^{-6}\text{N}$  zinc ions in the presence of  $10^{-3}\text{N}$  cadmium ions in 0.1N potassium chloride as supporting electrolyte (air-free). It should be pointed out that in contradistinction to conventional polarography, the small zinc wave can be suitably amplified for accurate measurement without interference from the preceding cadmium ion discharge.

As shown by equations (7), (8), and (9) in Part I of this series, the maximum differential current  $i_{D\sim}$  should increase with decreasing ohmic resistance of the cell circuit. This was found to be the case. The results for cadmium and zinc cited above referred to experiments carried out with the help of a saturated calomel reference electrode (resistance 1275 ohms) and a potentiometer with a resistance of 1000 ohms per volt. Using a low resistance potentiometer (22 ohms/volt)  $i_{D\sim}$  for cadmium and zinc increased appreciably. Still higher  $i_{D\sim}$  values were obtained with a mercury pool anode in place of the calomel reference electrode. Figure 3 shows clearly the improvements obtained by keeping the resistance of the cell circuit low. It can be seen that the calibration curves for cadmium and zinc approximate a straight line as expected from theoretical considerations (cf. Part I). The curve obtained by inserting a series resistance of 1500 ohms into the low resistance cell circuit was similar to that obtained with the saturated calomel electrode and the high resistance potentiometer.



In 0.1N hydrochloric acid (air-free), using a mercury pool anode, the summit potentials of cadmium and zinc ions were shifted by about 60 mV. to more negative values, due to the change in the reference potential. The  $E_s$  values in 0.1N hydrochloric acid remained constant over a tenfold dilution range. If, however, 0.1M acetate buffer was used as supporting electrolyte both the cadmium and zinc summit potentials were shifted to more positive values with increasing dilution.

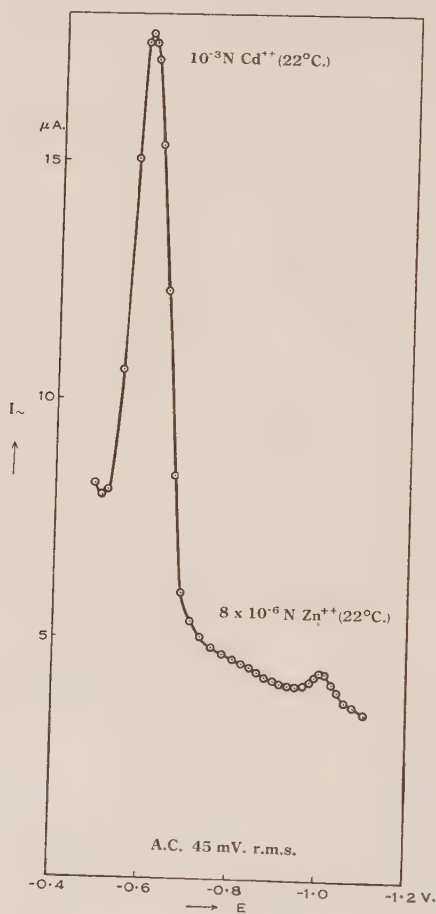
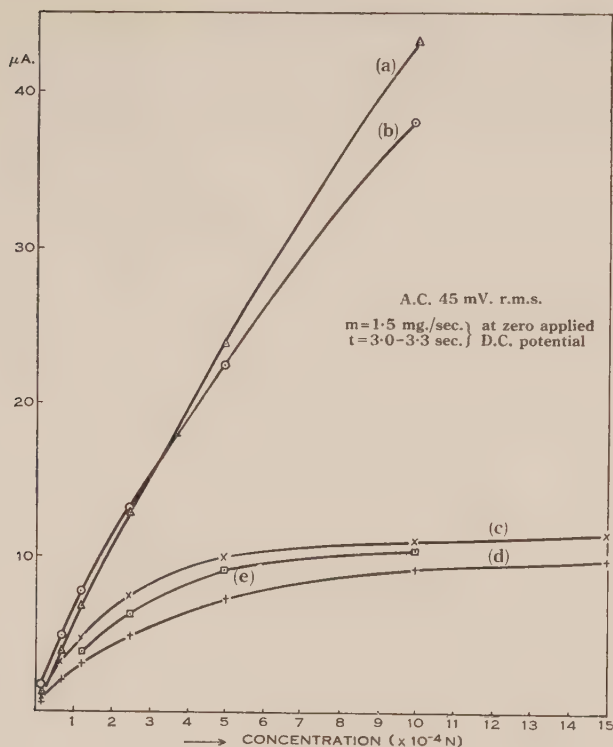


Fig. 2

The temperature coefficients of the  $i_{D\sim}$  values for cadmium and zinc were 0.4 per cent./°C. within the range of 20 to 40 °C.

The variation of wave height with the capillary distance from the pool anode was then investigated. No appreciable difference in the values of  $i_{D\sim}$  was observed when the distance of the capillary was varied from 20 to 10 mm. from the pool electrode (see Table 2).

Fig. 3.—Calibration curves of  $\text{Zn}^{++}$  and  $\text{Cd}^{++}$ .

- (a)  $\text{Zn}^{++}$  in 0.1N HCl (in air) at 20 °C.  $r$ , 170  $\Omega$  at zero applied D.C. potential. Potentiometer resistance, 22  $\Omega/\text{V}$ . Hg pool anode.
- (b)  $\text{Cd}^{++}$  in 0.1M acetate buffer, pH 4.7 (in air) at 20 °C.  $r$ , 225  $\Omega$  at zero applied potential. Potentiometer resistance, 22  $\Omega/\text{V}$ . Hg pool anode.
- (c)  $\text{Cd}^{++}$  in 0.1M acetate buffer, pH 4.7 (in air) at 25 °C.  $r$ , 1385  $\Omega$  at zero applied potential. Potentiometer resistance, 1000  $\Omega/\text{V}$ . S.C.E. reference electrode.
- (d)  $\text{Zn}^{++}$  as in (c).
- (e)  $\text{Cd}^{++}$  as in (b), 1500  $\Omega$  in series.

TABLE 2

VARIATION OF WAVE HEIGHT WITH CAPILLARY DISTANCE  
 Solution:  $10^{-3}\text{N}$   $\text{Cd}^{++}$  in 0.1M acetate buffer (in air).  
 Temperature, 20 °C. Surface area of anode, 8.5  $\text{cm}^2$

Distance of Capillary from Anode Surface (mm.)	$i_D \sim$ ( $\mu\text{A.}$ )
3	33
5	32
10	30.5
15	30.3
20	30.0
30	29.8
40	28.2

Experiments were carried out to observe whether the presence of other dischargeable ions had any influence on the maximum differential currents of cadmium and zinc. Results showed that the presence of nobler ion species had no effect on the  $i_{D\sim}$  values obtained as shown by the results reported in Table 3.

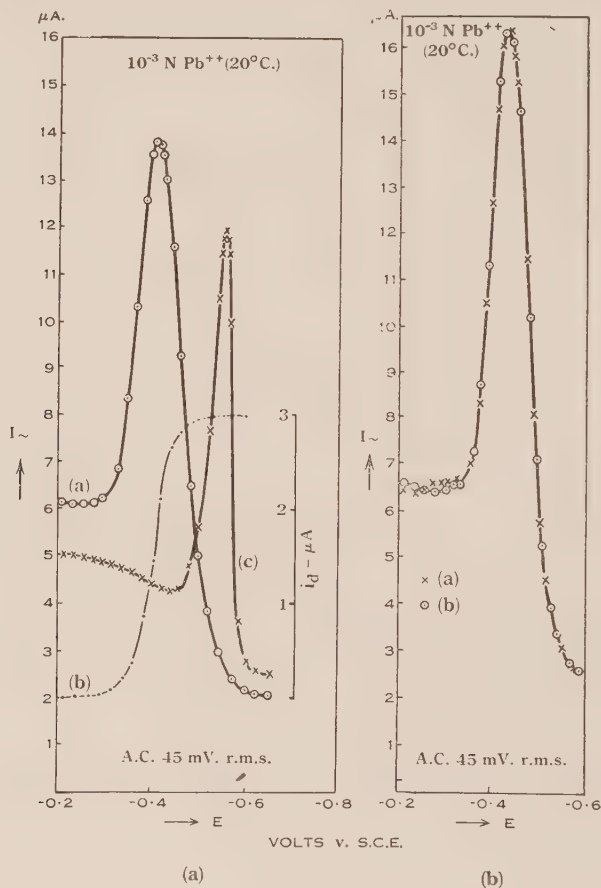


Fig. 4.—A.C. polarogram of  $10^{-3}N$   $Pb^{++}$  at  $20^{\circ}C$ . Capillary characteristics:  $m$ , 1.5 mg./sec.  $t$ , 3.0 sec. at zero applied potential. Potentiometer, 22  $\Omega/V$ .  $r$ , 1300  $\Omega$ . S.C.E. reference electrode.

- (a) (a) 0.1N KCl supporting electrolyte (air-free).
- (b) D.C. polarogram of (a).
- (c) 0.1N KCl supporting electrolyte (in air).
- (b) (a) 0.1M acetate buffer supporting electrolyte (in air).
- (b) 0.1M acetate buffer supporting electrolyte (air-free).

It can therefore be concluded that the presence of previously discharged ions in the solution does not affect the magnitude of the differential current of the ions under consideration.

TABLE 3

INFLUENCE OF DISCHARGEABLE IONS ON THE MAXIMUM DIFFERENTIAL CURRENTS OF Cd AND Zn

Solution Analysed	Mixture of Nobler Ion Species Added ( $10^{-3}\text{N}$ )	$i_{D\sim}$ before Addition ( $\mu\text{A.}$ )	$i_{D\sim}$ after Addition ( $\mu\text{A.}$ )	Difference (%)
$10^{-3}\text{N Cd}^{++}$ in $0.5\text{N HCl}$ (in air)	$\text{Bi}^{+++}$ , $\text{Sb}^{+++}$ , $\text{Pb}^{++}$ , $\text{Tl}^{+}$	31.5	31.6	+0.3
$10^{-3}\text{N Zn}^{++}$ in $0.1\text{N HCl}$ (in air)	$\text{Bi}^{+++}$ , $\text{Sb}^{+++}$ , $\text{Tl}^{+}$ , $\text{Pb}^{++}$ , $\text{In}^{+++}$ , $\text{Cd}^{++}$	43.4	43.1	-0.7

## III. LEAD

The A.C. polarographic waves of lead in the presence and absence of oxygen, and using different supporting electrolytes, are shown in Figures 4 (a) and 4 (b). The  $E_s$  value of polarogram (a) in Figure 4 (a) ( $-0.40\text{ V. v. S.C.E.}$ ) corresponds to the reported D.C. half-step potential(1). In  $0.1\text{N}$  potassium chloride solution the summit potential shifted  $150\text{ mV.}$  to more negative values in the presence of air. At the same time a lowering of the wave height was observed (cf. Fig. 4 (a), polarogram (c)). These changes were absent when  $0.1\text{N}$  hydrochloric acid was used as supporting electrolyte. It can, therefore, be concluded that  $\text{Pb}^{++}$  similarly to  $\text{Zn}^{++}$  is transformed into a hydroxo-complex when polarographed in air. In contradistinction to  $\text{Zn}^{++}$ , however, an A.C. polarogram is obtained here. This is explained by the fact that the biplumbite ion is reversibly reduced at the dropping electrode (3) whereas the zincate ion is not.

## IV. THALLIUM

Thallous ion gave well-defined A.C. waves both in the presence and absence of air and in either  $0.1\text{N}$  potassium chloride or  $0.1\text{M}$  acetate buffer as supporting electrolyte. The summit potential in either solution was found to be  $-0.47\text{ V. (v. S.C.E.)}$  in accordance with the D.C. half-step potential(4-6). The summit potential was unaffected by the presence of air.

Lead and thallium were polarographically separable in  $0.1\text{N}$  potassium chloride as supporting electrolyte. Their corresponding summit potentials in air are about  $50\text{ mV.}$  apart, sufficient for separation in A.C. polarography. In the presence of nitrogen, however, the lead and thallium waves nearly coalesced as shown in Figure 5.

## V. INDIUM

Indium gave sharp, well-defined A.C. waves in either potassium chloride ( $0.1$  or  $0.5\text{N}$ ) or hydrochloric acid ( $0.1$  or  $0.5\text{N}$ ) as well as in  $0.1\text{M}$  acetate buffer ( $\text{pH } 4.7$ ) as supporting electrolyte. In potassium chloride or acetate



buffer, however, a slight hydrolysis of the indium salt occurred after allowing the solution to stand for some time.

The summit potentials were unaffected by the presence of air. They corresponded to the half-step potentials as reported ( $-0.565$  V. in  $0.1N$  KCl *v. S.C.E.*, 5, 7).

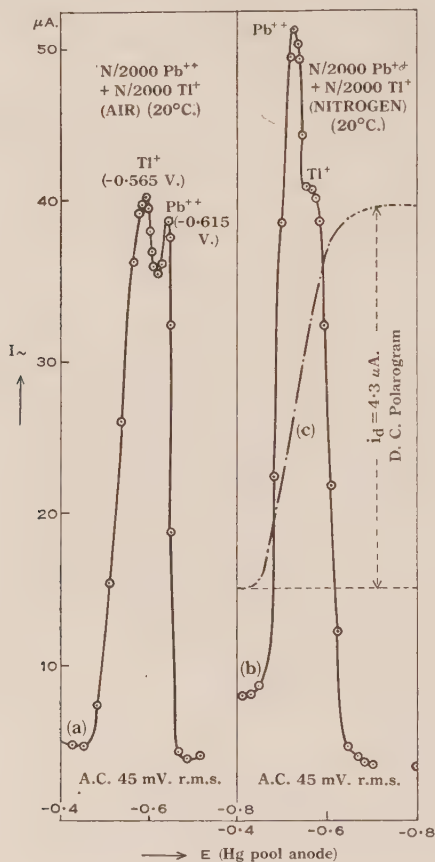


Fig. 5

## VI. BISMUTH AND ANTIMONY

Both bismuth and antimony ions are reversibly reduced at the dropping electrode(4). D.C. polarography in  $0.1N$  hydrochloric acid solution, however, showed maxima when the  $Sb^{+++}$  or  $Bi^{+++}$  concentration was  $10^{-3}N$  or higher. In  $0.5N$  hydrochloric acid, and at concentrations less than  $10^{-3}N$  of  $Bi^{+++}$  and  $Sb^{+++}$ , the maxima were greatly reduced and good A.C. polarographic waves were obtained. The respective  $E_s$  values found were  $-0.135$  V. for  $Sb^{+++}$  and  $-0.065$  V. for  $Bi^{+++}$  (both *v. S.C.E.*).

## VII. SEPARABILITY OF A.C. POLAROGRAPHIC WAVES

In  $0.1N$  hydrochloric acid and air the summit potentials of indium and cadmium were about  $-0.565$  and  $-0.610$  V. *v. S.C.E.* respectively. These

values correspond closely to the respective half-wave potentials as obtained experimentally under free-air conditions. When a low A.C. input voltage was used (5 to 15 mV. r.m.s.) the indium and cadmium waves were clearly separable. Coalescence of cadmium and indium waves occurred, however, when the superposed A.C. potential exceeded the difference in the  $E_s$  potentials of the two ion species. Better resolution of the waves was obtained when a low impedance cell was used (i.e. mercury pool anode). Figure 6 shows the sharp separation

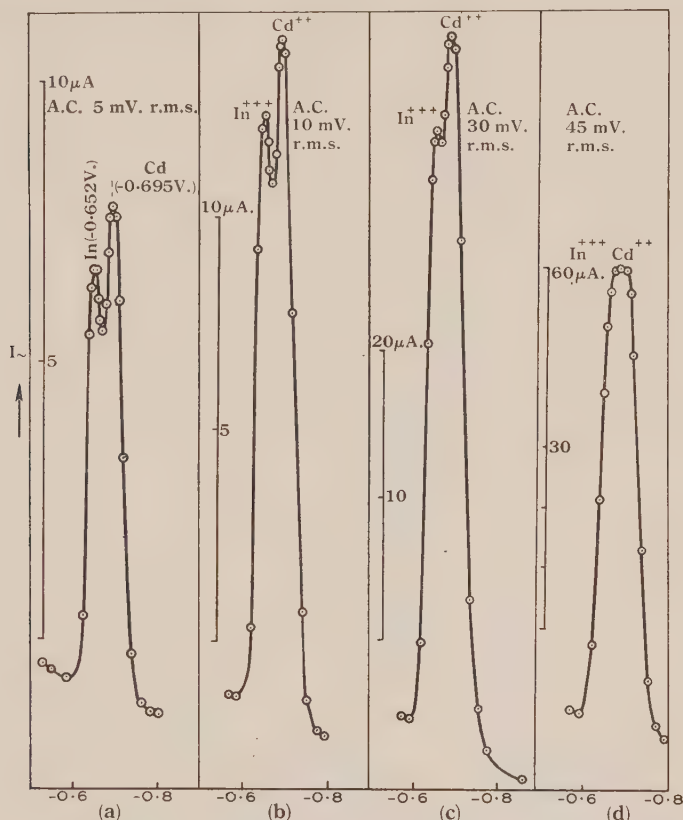


Fig. 6.—Separability of  $1.33 \times 10^{-3} \text{N In}^{+++} + 10^{-3} \text{N Cd}^{++}$  in  $0.1 \text{N HCl}$  at  $20^\circ \text{C}$ . (in air). Hg pool anode. Capillary characteristics:  $m$ ,  $1.5 \text{ mg./sec}$ .  $t$ ,  $3.1 \text{ sec}$ . at zero applied potential. Potentiometer,  $22 \Omega/\text{V}$ .

of the peaks of the indium and cadmium waves using 5 and 10 mV. A.C. respectively. No separation, however, was obtained with 45 mV. A.C. and over.

Figure 7 contrasts the A.C. and D.C. polarograms of  $10^{-3} \text{N}$  lead, thallium, indium, cadmium, and zinc in  $0.1 \text{N}$  hydrochloric acid. The zinc wave is considerably suppressed, which might be due either to the beginning of  $\text{H}^+$  discharge or to complex formation. In conventional polarography the limit of separation is 60 mV. This is just sufficient to see that there are two inflections

in the polarogram but no accurate measurements can be made(8). By the use of A.C. polarography, however, indium and cadmium whose summit potentials were found to be only 40 mV. apart ( $-0.645$  V. for  $\text{In}^{+++}$  and  $-0.685$  V. for  $\text{Cd}^{++}$ , pool anode) were clearly separable.

Two ion species whose A.C. waves overlap can be quantitatively analysed by the following procedure: Let  $C$  be the concentration of two ion species  $i$  and  $j$  in solution. The respective summit potentials ( $E_{si}$  and  $E_{sj}$ ) are  $\Delta E_s$

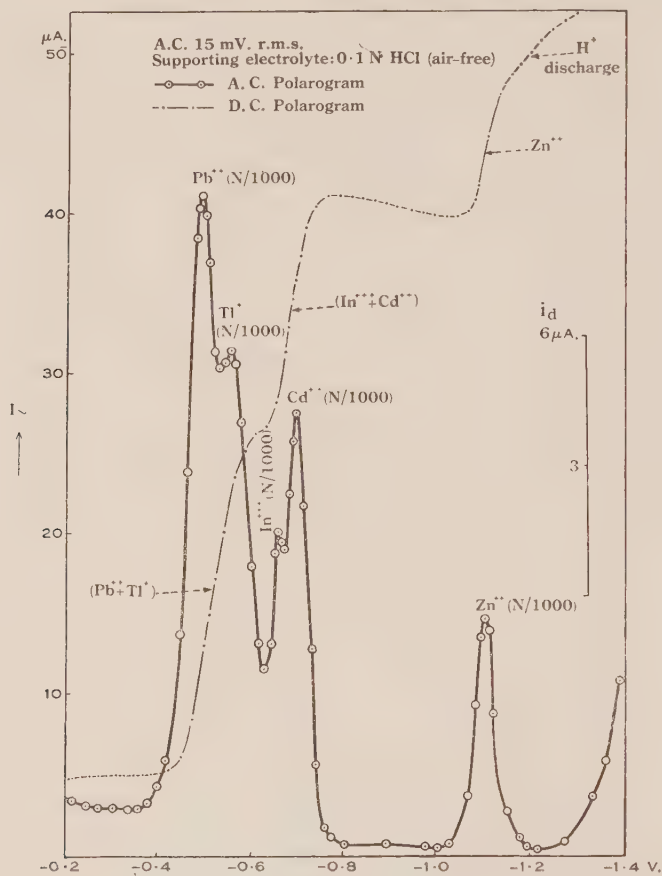


Fig. 7

volts apart as shown in Figure 8.  $OAQ$  represents the A.C. wave of ion species  $i$  and  $PBR$  that of ion species  $j$  respectively. The wave  $OXYR$  represents the combined polarogram of  $i$  and  $j$ , i.e. the addition curve of  $OAQ$  and  $PBR$ . Thus

$$XL = AL + PL$$

and

$$YM = BM + QM.$$

$AL$  is the magnitude of the maximum differential current of species  $i$  and  $BM$  that of species  $j$  at their respective summit potentials.  $PL$  represents the differential current of ion species  $j$  at the applied D.C. potential  $E_{si}$ , whereas

$QM$  represents the differential current of ion species  $i$  at the D.C. potential  $E_{sj}$ .  $OLMR$  is the base current of the A.C. polarogram. It is thus possible to calculate  $XL$  and  $YM$  respectively from the ratios  $QM/AL$  and  $PL/BM$ . Conversely it is possible to calculate  $AL$  and  $BM$ , knowing  $XL$  and  $YM$ .

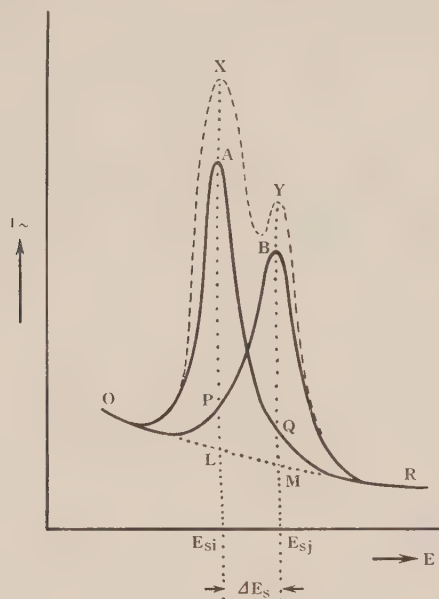


Fig. 8

TABLE 4  
QUANTITATIVE ANALYSIS OF OVERLAPPING WAVES

Experiment No.	Ion	Concentration	Supporting Electrolyte	$i_{D_1\sim}$	$i_{D_2\sim}$	$i_{D_3\sim}$	Difference between $i_{D_1\sim}$ and $i_{D_3\sim}$ (%)
				$i_{D\sim}$ from Separate Polarograms ( $\mu A.$ )	$i_{D\sim}$ from Combined Polarograms ( $\mu A.$ )	$i_{D\sim}$ Corrected from Combined Polarograms ( $\mu A.$ )	
1	In <sup>+++</sup>	$10^{-3}N$	0.5N HCl	20.3	24.6	19.6	-3.5
	Cd <sup>++</sup>	$10^{-3}N$	,,	30.5	31.1	30.2	-1.0
2	In <sup>+++</sup>	$5 \times 10^{-4}N$	,,	10.7	11.9	10.4	-3.0
	Cd <sup>++</sup>	$2.5 \times 10^{-4}N$	,,	7.85	8.2	7.5	-4.5
3	In <sup>+++</sup>	$5 \times 10^{-4}N$	,,	10.7	12.8	10.2	-5.0
	Cd <sup>++</sup>	$5 \times 10^{-4}N$	,,	16.2	16.5	15.6	-3.7



This procedure was adopted when analysing a solution containing both cadmium and indium ions. The A.C. waves of  $10^{-3}N$  cadmium and indium were first recorded separately and the respective values corresponding to *AL*, *PL*, *BM*, and *QM* (Fig. 8) observed. The polarogram of a  $10^{-3}N$  cadmium +  $10^{-3}N$  indium solution was then taken and the apparent values of  $i_{D\sim}$  for cadmium (*YM*) and indium (*XL*) noted. From these the true values of  $i_{D\sim}$  were calculated. The results are reported in Table 4.

### VIII. SIMULTANEOUS A.C. POLAROGRAPHY OF BISMUTH, ANTIMONY, LEAD, THALLIUM, INDIUM, CADMIUM, AND ZINC

The A.C. polarographic "spectrum" of a solution containing all the seven elements discussed above is shown in Figure 9.

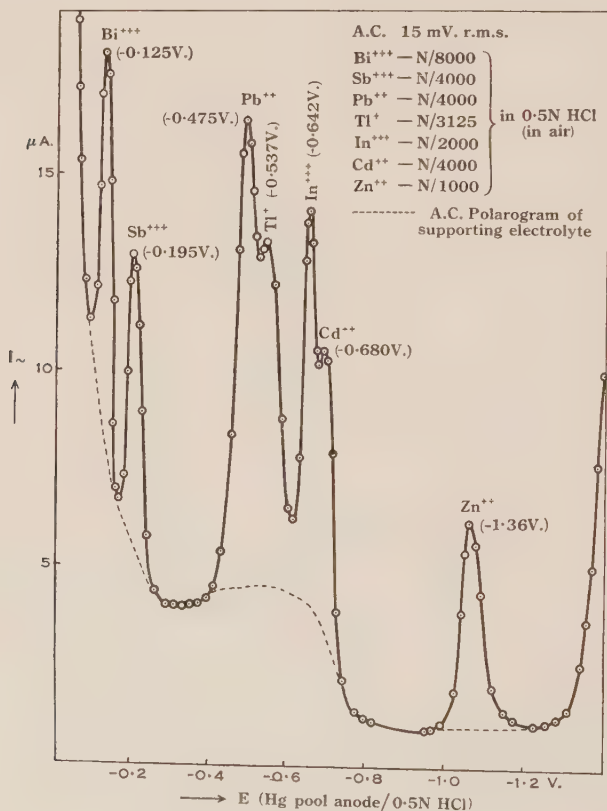


Fig. 9

The respective calibration curves are given in Figure 10 showing the very nearly linear relationship between the logarithm of the maximum differential current and the logarithm of the ion-concentration over a wide dilution range.

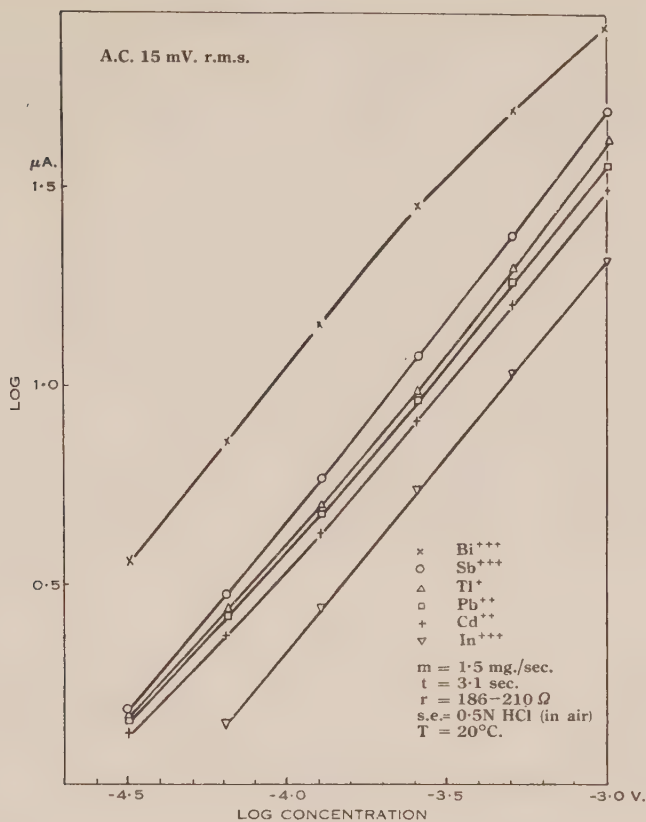


Fig. 10

TABLE 5

VALUES OF DIFFERENTIAL AND D.C. DIFFUSION CURRENTS OF IONS INVESTIGATED

Ion Species	Concentration	Temperature (°C.)	Supporting Electrolyte	$m$ (mg./sec.)	$t$ (sec.)	$r$ (ohms)	$i_d$ (μA.)	$i_{D\sim}$ (μA.)
Bi+++	$10^{-3}N$	20	0.5N HCl	1.5	3.2	173	3.25	74.3
Sb+++	"	"	"	"	3.2	174	3.50	46.5
Pb++	"	"	"	"	3.3	181	2.70	35.0
Tl+	"	"	"	"	3.2	183	3.85	40.0
In+++	"	"	"	"	3.1	185	2.75	20.0
Cd++	"	"	"	"	3.0	186	2.65	30.0
Zn++	"	"	0.1N HCl	"	2.9	195	2.65	16.2

Table 5 compares the values of the differential currents ( $i_{D\sim}$ ) and the D.C. diffusion currents ( $i_d$ ) of the ions investigated.

## IX. ACKNOWLEDGMENTS

The authors are indebted to the Trustees of the Science and Industry Endowment Fund for a grant which made this work possible. Thanks are also due to Dr. F. P. J. Dwyer, University of Sydney, for the supply of a sample of pure indium metal.

## X. REFERENCES

- (1) KOLTHOFF, I. M., and LINGANE, J. J.—“Polarography.” 1st Ed. p. 482. (Interscience Publishers Inc.: New York, 1941.)
- (2) KOLTHOFF, I. M.—*Industr. Engng. Chem. (Anal. Ed.)* **14**: 195 (1942).
- (3) KOLTHOFF, I. M., and LINGANE, J. J.—“Polarography.” 1st Ed. p. 267. (Interscience Publishers Inc.: New York, 1941.)
- (4) HEYROVSKY, J., and ILKOVIC, D.—*Coll. Trav. Chim. Tchécosl.* **7**: 198 (1935).
- (5) LINGANE, J. J.—*J. Amer. Chem. Soc.* **61**: 2099 (1939).
- (6) KLUMPAR, J.—*Coll. Trav. Chim. Tchécosl.* **11**: 459 (1939).
- (7) KOLTHOFF, I. M., and LINGANE, J. J.—*Chem. Rev.* **24**: 1 (1939).
- (8) MÜLLER, O. H.—*J. Chem. Educ.* **18**: 322 (1941).

# THE CRYSTAL STRUCTURE OF INDIUM MONOBROMIDE

By N. C. STEPHENSON\* and D. P. MELLOR\*

[*Manuscript received June 13, 1950*]

## Summary

The crystal structure of indium monobromide has been determined using powder and complete rotation photographs. The unit cell dimensions are

$$a = 4.46 \pm 0.005 \text{ \AA}$$

$$b = 12.39 \pm 0.02 \text{ \AA}$$

$$c = 4.73 \pm 0.01 \text{ \AA}$$

with four molecules per cell. The space group is  $D_{2h}^{17}$ -*Cmcm*. The structure is a layer structure isomorphous with that of thallium iodide TlI. Each indium has five bromine atoms arranged about it at the corners of a rectangular pyramid with one In-Br bond of 2.80 Å and four In-Br bonds of 3.29 Å. Indium atoms are similarly arranged about bromine atoms.

## I. INTRODUCTION

Compounds of the type  $AB$ , in which  $A$  and  $B$  are ions with rare gas electron configurations possess relatively simple, close-packed structures whose nature is largely determined by the radius ratio of the two ions. On the other hand,  $AB$  compounds in which  $A$  has an unshared electron pair exhibit more complicated structures such as those of PbO, SnS, and TlI. Since InBr falls into this latter class it was thought to be of some interest to determine its structure.

## II. PREPARATION OF InBr

Direct bromination of molten indium with excess bromine yields white, strongly deliquescent crystals of InBr<sub>3</sub>. These can be readily reduced at about 200 °C. with excess indium to yield an orange-red, polycrystalline mass of InBr. Though sensitive to moisture, which causes disproportionation to In and InBr<sub>3</sub>, the monobromide can be handled in air with reasonable precautions. After powder photographs had shown that the symmetry of crystals of InBr was lower than that of the cubic, hexagonal, or tetragonal systems, it was decided to attempt the growing of single crystals. Two methods were tried:

- (a) Growth from the melt.
- (b) Growth from the vapour phase.

Both the generally used methods for (a) were tried but without success. Vacuum sublimation yielded small flakes which, although without well-defined crystal faces except those parallel to the flake surfaces, proved to be individual crystals. Under the polarizing microscope fragments of the flakes were doubly refracting and strongly pleochroic.

\* Department of Chemistry, University of Sydney.



## III. DETERMINATION OF UNIT CELL

Three complete rotation photographs were taken, the first being about a direction perpendicular to a crystal flake. A spacing of  $12.4 \text{ \AA}$  was found in this direction. In order to orient a flake for the other photographs, a Laue photograph was taken using radiation from a copper target with the beam

TABLE 1  
POWDER LINES OF InBr

<i>hkl</i>	Experimental $\sin^2\theta$	Calculated $\sin^2\theta$	Intensity
110	0.0338	0.0337	<i>w</i>
021	0.0418	0.0421	<i>vw</i>
111	0.0599	0.06029	<i>vs</i>
040	0.0620	0.0620	<i>s</i>
041	0.0886	0.0886	<i>vw</i>
131	0.0918	0.0913	<i>s</i>
002	0.1055	0.1065	<i>ms</i>
200	0.1188	0.1192	<i>w</i>
150	0.1267	0.1268	<i>ms</i>
151	0.1533	0.1534	<i>ms</i>
042	0.1682	0.1684	<i>ms</i>
240	0.1808	0.1811	<i>ms</i>
241	0.2075	0.2077	<i>vvw</i>
202	0.2250	0.2251	<i>vvw</i>
152	0.2326	0.2333	<i>w</i>
242	0.2874	0.2876	<i>w</i>
190	0.3444	0.3433	<i>vvw</i>
281	0.3935	0.3936	<i>vw</i>
243	0.4204	0.4207	<i>vvw</i>
192	0.4501	0.4498	<i>vvw</i>
282	0.4730	0.4735	<i>vvw</i>

perpendicular to the flake. On this photograph, which showed the presence of two mirror planes, there were four strong spots at the corners of what was almost a square. These spots were used to orient the crystal by a process of trial and error. Two directions of rotation were fixed so that they were parallel to the edges of the rectangle formed by the four spots. From these photographs the spacings along the two axes were found to be  $4.4$  and  $4.7 \text{ \AA}$  respectively. A unit cell with spacings  $12.4$ ,  $4.7$ , and  $4.4 \text{ \AA}$  was provisionally adopted. Since all the rotation photographs showed two mutually perpendicular mirror planes it was assumed that the crystal possessed orthorhombic symmetry. On the basis of the above unit cell it was found possible to index (see Table 1) all the lines of a power photograph taken with a  $19 \text{ cm.}$  camera. In this way the cell dimensions were obtained more accurately. The unit cell finally adopted was

$$a = 4.46 \pm 0.005 \text{ \AA}$$

$$b = 12.39 \pm 0.02 \text{ \AA}$$

$$c = 4.73 \pm 0.01 \text{ \AA}.$$

From the reported(1) density of the crystals (4.96) the calculated number of molecules per unit cell is four.

#### IV. SPACE GROUP DETERMINATION

After indexing the rotation photographs with the aid of Bernal charts the following regularities were noted :

- (a)  $hkl$  planes were present only when  $h+k=2n$   
 (b)  $okl$     "    "    "    "    "     $k=2n$   
 (c)  $hol$     "    "    "    "    "     $h=2n$  and  $l=2n$ .

These absences are demanded by the three space groups  $D_{2h}^{17}$ ,  $C_{2v}^{12}$ , and  $C_{2v}^{16}$ . Of these space groups  $D_{2h}^{17}$  is the only one that has a centre of symmetry. Application of Martin's test for pyroelectricity showed that there was no evidence for the absence of a centre of symmetry in the crystals of InBr. On the basis of this negative evidence  $D_{2h}^{17}$  was provisionally adopted as the space group. Agreement between the calculated and observed intensities of diffraction spots justified the adoption of this space group. Four atoms can be placed in the unit cell in the following way for  $D_{2h}^{17}$ :  $o, y, \frac{1}{4}$ ;  $o, y, \frac{3}{4}$ ;  $\frac{1}{2}, y+\frac{1}{2}, \frac{1}{4}$ ;  $\frac{1}{2}, \frac{1}{2}-y, \frac{3}{4}$ .

Fixing atomic positions therefore involves finding two parameters:  $y$  for In, and  $y'$  for Br.

#### V. CALCULATION OF PARAMETERS

Since (002) is the most intense of all the (001) reflections it is evident that the atom with the larger scattering power (In) must have the larger of the two parameters. If a bromine atom of radius 1.19 Å is to be placed in the unit cell of InBr, the parameter must, from considerations of space available for it, be greater than 0.1. Similarly if the atomic radius of In is taken as 1.45 Å its parameter must be greater than 0.2. The parameters were found by a method of trial and error based essentially on the inequalities of intensities of diffraction spots observed on complete rotation photographs.

Intensities were calculated from the formula

$$\sqrt{I} = \sqrt{l p m F} \varphi,$$

where  $\varphi$ =multiplicity of the reflection obtained from the reciprocal lattice diagram and cell symmetry,

$m$ =number of equivalent atoms in the cell,

$F$ =the simplified structure factor,

$l$  and  $p$  are the Lorentz and polarization factors respectively(2). The intensities of the observed reflections from the planes (002), (004), (200), (202), (204), (400) which are independent of the parameters  $y$  and  $y'$  were used to provide a qualitative scale on which to compare the observed intensities of all the other planes. In all a series consisting of 26 reflections observed on a complete rotation photograph were studied and listed in their order of magnitude. By varying the bromine parameter  $y'$  from 0.10 to 0.16 in intervals of 0.01 and the indium parameter  $y$  from 0.21 to 0.40 in the same intervals, the value of  $\sqrt{I}$  was calculated for every possible pair or parameters for every reflection. The

calculated intensities were entered at the corresponding points on a large graph whose coordinate axes were the two parameters. The parameters were fixed by choosing that point on the graph which gave the best agreement between calculated and observed intensities. In this manner the parameters were narrowed down to

$$y = 0.386,$$

$$y' = 0.160.$$

A set of calculated and observed intensities is given in Table 2.

TABLE 2  
CALCULATED AND OBSERVED INTENSITIES

<i>hkl</i>	Observed	Calculated	<i>hkl</i>	Observed	Calculated
002	<i>vs</i>	416	192	<i>s</i>	386
004	<i>w</i>	186	200	<i>vs</i>	391
021	<i>s</i>	260	202	<i>vs</i>	267
023	<i>vw</i>	114	204	<i>w</i>	171
040	<i>s</i>	435	221	<i>w</i>	148
041	<i>vs</i>	432	223	<i>vvw</i>	90
042	<i>vs</i>	535	240	<i>ms</i>	258
043	<i>w</i>	213	241	<i>ms</i>	270
044	<i>w</i>	290	242	<i>ms</i>	507
080	<i>w</i>	111	243	<i>w</i>	167
081	<i>ms</i>	340	244	<i>ms</i>	270
082	<i>vw</i>	174	280	<i>vvw</i>	63
110	<i>s</i>	298	281	<i>w</i>	260
111	<i>vs</i>	789	282	<i>vw</i>	162
113	<i>s</i>	354	310	<i>vw</i>	112
131	<i>vs</i>	494	311	<i>s</i>	334
151	<i>w</i>	266	331	<i>w</i>	237
152	<i>ms</i>	374	350	<i>vw</i>	132
153	<i>ms</i>	255	351	<i>w</i>	246
171	<i>w</i>	152	400	<i>w</i>	178
192	<i>s</i>	386	440	<i>vw</i>	145

The parameters lead to the following interatomic distances :

$$4\text{In-Br} \quad 3.29 \text{ \AA}$$

$$1\text{In-Br} \quad 2.80 \text{ \AA}$$

$$2\text{Br-Br} \quad 4.61 \text{ \AA}$$

$$2\text{In-In} \quad 3.68 \text{ \AA}$$

These distances are shown in Figure 1.

## VI. DESCRIPTION OF STRUCTURE

In drawing Figure 1 the origin that was used for calculation of intensities has been shifted  $c/4$  so that some of the atoms in the lower half of the cell lie on the cell edges.

Except for a difference in parameters, the structure of InBr is the same as that of TlI(3). Each indium atom has five bromine atoms arranged about it at the vertices of a rectangular pyramid with one In-Br bond of 2.80 Å and four In-Br bonds 3.29 Å. The indium atom is elevated a distance 0.57 Å from the plane of the four bromine atoms that form the base of the rectangular pyramid. Indium atoms are similarly arranged about each bromine. The resulting arrangement is a layer structure, one layer of which is shown in Plate 1.

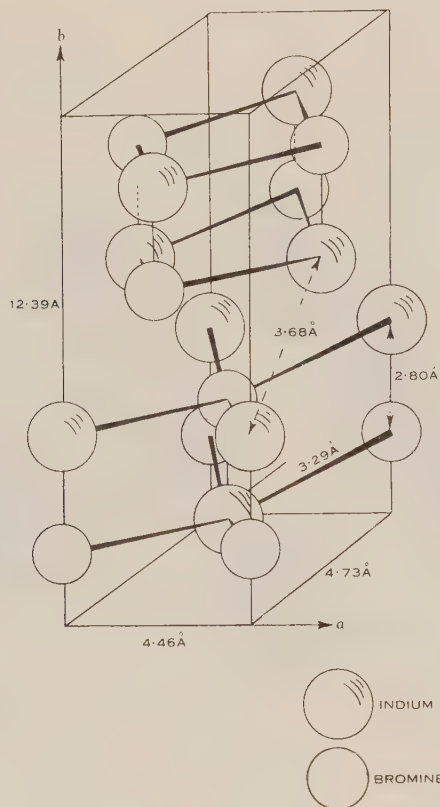


Fig. 1.—The arrangement of atoms in the unit cell of InBr.

The separation distance between this layer and the identical but slightly displaced ( $c/2$ ) layer that fits on top of it is determined by In-In contacts. This gives the van der Waals radius (ionic) of  $\text{In}^{\text{I}}$  as 1.84 Å. If the covalent radius of bromine is taken as 1.19 Å the shortest In-Br distance leads to a covalent radius of 1.61 Å for  $\text{In}^{\text{I}}$ .

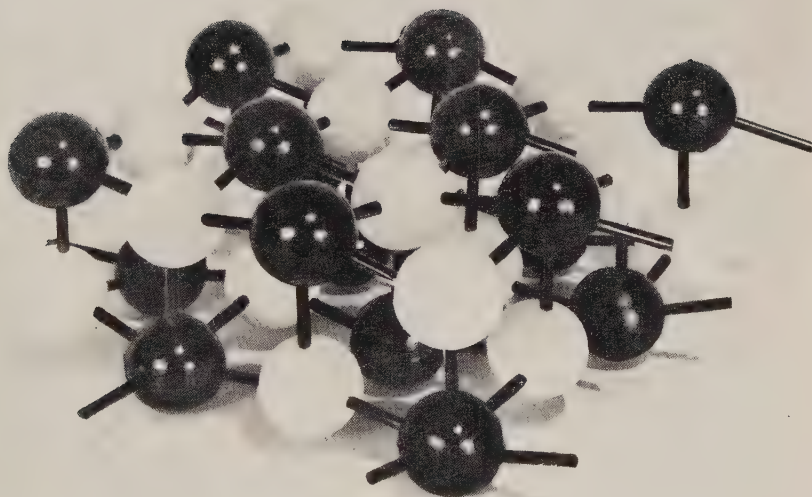
#### VII. ACKNOWLEDGMENT

We are indebted to the Commonwealth Research Fund for providing some of the equipment used in this work.



## VIII. REFERENCES

- (1) HODGMAN, C. S. (Ed.).—"Handbook of Chemistry and Physics." 28th Ed. p. 390.  
(Chemical Publishing Co. : Cleveland, 1944.)
- (2) "Numerical Structures Factor Tables." (M. J. Buerger).—*Geol. Soc. Amer. Spec. Pap.*  
33, 1941.
- (3) HELMHOLTZ, L.—*Z. Krystallogr.* 95 : 129 (1936).



Model to illustrate the layer structure of indium monobromide.



# THE IODINATION OF AROMATIC COMPOUNDS

## IV. THE IODINATION OF AROMATIC HYDROCARBONS AND NUCLEAR-SUBSTITUTED PHENOLS

By L. JURD\*

[*Manuscript received May 25, 1950*]

### *Summary*

Salicylic acid and *p*-nitrophenol are almost quantitatively diiodinated when treated in alcoholic solution with iodine, hydrogen peroxide, and sulphuric, nitric, hydrochloric, or phosphoric acid. In glacial acetic acid solution salicylic acid gives good yields of the diiodo compound only in the presence of hydrochloric acid. No appreciable iodination occurs in the presence of sulphuric or nitric acid whilst 3,5-dibromosalicylic acid is formed with hydrobromic acid.

Cinnamic acid reacts slowly with iodine, hydrogen peroxide, and sulphuric acid in methanol and ethanol solutions to give the corresponding  $\alpha$ -iodo- $\beta$ -alkoxy- $\beta$ -phenyl-propionic acids.

Benzene homologues are readily iodinated in alcoholic solution in the presence of sulphuric or nitric acid although benzene does not react appreciably under similar conditions.

### I. INTRODUCTION

Of the many methods proposed for the iodination of phenols containing nuclear-deactivating substituents, that employing iodine monochloride is by far the most satisfactory. Other reagents, for example, iodine and alkali, ammonia, iodic acid, or mercuric oxide, apparently give poor yields of impure products. These various procedures were examined by Woollett and Johnson(1) and Block and Powell(2) who obtained 3,5-diiodosalicylic acid and 2,6-diiodo-4-nitrophenol respectively in 85–90 per cent. yields by the action of iodine monochloride on salicylic acid and *p*-nitrophenol.

Iodohydrocarbons, on the other hand, are prepared most conveniently from the corresponding amines through the diazo reaction(3), although aromatic hydrocarbons do react directly with iodine and an oxidizing agent to give nuclear-iodinated derivatives. Concentrated nitric acid(4), sodium persulphate(5), and iodic acid(6) are the most effective oxidizing agents, toluene and the xylenes giving 60–70 per cent. yields of the monoiodo derivatives.

In 1927 Marsh(7) reported the iodination of phenol and 2-naphthol in alcoholic solution by treatment with iodine and hydrogen peroxide. It was shown recently, however, that, although the relatively non-reactive phenolic ethers are not appreciably iodinated by Marsh's process, the addition of a concentrated mineral acid to the alcoholic mixture results in good yields of the

\* Massey Agricultural College, Palmerston North, New Zealand. Present address : Department of Chemistry, University of Nottingham, England.



iodo-ethers(8). The investigation of this reaction has now been extended, the influence of mineral acids on the iodination of aromatic hydrocarbons and less easily substituted phenols in alcoholic and acetic acid solutions being examined.

## II. THE INFLUENCE OF MINERAL ACIDS AND THE SOLVENT ON THE IODINATION OF PHENOLS AND HYDROCARBONS

It has now been found that equimolecular quantities of iodine and salicylic acid react slowly in ethanol solution in the presence of hydrogen peroxide to give 3,5-diiodosalicylic acid. Addition of concentrated sulphuric, nitric, or phosphoric acid to the alcoholic reaction mixture accelerates the iodination very considerably and almost quantitative yields of the diiodo compound result. Hydrochloric acid gives somewhat smaller yields of a less pure product.

Although hydrogen peroxide and equimolecular quantities of iodine and salicylic acid react slowly in glacial acetic acid solution, no appreciable iodination occurs within several hours when sulphuric or nitric acid is added. Hydrochloric acid, on the other hand, gives 75 per cent. yields of the diiodo-phenol within 15 minutes. In the presence of hydrobromic acid 3,5-dibromosalicylic acid is the chief product, although some iodination does occur. When iodine is absent chlorination of salicylic acid occurs in the presence of hydrochloric acid(9) and bromination in the presence of hydrobromic acid.

The products obtained in the halogenation of salicylic acid under various conditions are summarized in Table 1.

Like salicylic acid, *p*-nitrophenol reacts with an equimolecular quantity of iodine in alcoholic solution in the presence of sulphuric acid and hydrogen peroxide to give almost quantitative yields of 2,6-diiodo-4-nitrophenol.

The direct moniodination of salicylic acid and *p*-nitrophenol by treatment with iodine (0.5 mol.), hydrogen peroxide, and mineral acid is impracticable since considerable diiodination occurs and the similar solubilities of the products render their separation difficult. Thus 5-iodosalicylic acid, in 25 per cent. yields, and 3,5-diiodosalicylic acid, which can be separated partially by means of their barium salts, are obtained when salicylic acid is iodinated in a glacial acetic acid, hydrochloric acid mixture.

Cinnamic acid reacts slowly with iodine, hydrogen peroxide, and sulphuric acid in methanol and ethanol solutions to give  $\alpha$ -iodo- $\beta$ -methoxy- and  $\alpha$ -iodo- $\beta$ -ethoxy- $\beta$ -phenylpropionic acid respectively. Jackson and Pasiut(10) have prepared these iodo-alkoxy acids by the action of methyl and ethyl alcohol on  $\alpha$ -iodo- $\beta$ -chloro- $\beta$ -phenylpropionic acid.

Although benzene is not appreciably iodinated, the homologues of benzene react readily in ethanol solution with 0.5 molecular equivalents of iodine, hydrogen peroxide, and sulphuric or nitric acid to give good yields of the iodo derivatives. When hydrochloric acid is employed under similar conditions no iodination occurs.

Because of its simplicity and the high yields obtained this process of iodination appears to be preferable in many respects to that employing iodine monochloride. The method is capable of wide application for it has been found

TABLE 1  
THE INFLUENCE OF MINERAL ACIDS AND THE SOLVENT ON THE IODINATION OF SALICYLIC ACID IN THE PRESENCE OF HYDROGEN PEROXIDE  
Salicylic Acid (2.0 g.); Iodine (3.7 g.)

Solvent	Mineral Acid	Reaction Temperature (°C.)	Reaction Period (min.)	30% Hydrogen Peroxide (c.c.)	Product	Solvent Recrystallized	Yield (a) (%)	Melting Point (b) (°C.)
Ethanol (95%) (10.0 cc.)	Not added	70-75	75	3.5	3,5-Diiodosalicylic acid	Aqueous acetone	87	226
	H <sub>2</sub> SO <sub>4</sub> (1.0 cc.; sp. gr. 1.83)	60	15	2.3	"	"	91	234
	HNO <sub>3</sub> (1.0 cc.; sp. gr. 1.42)	65-70	15	2.5	"	"	94	228
	HCl (2.0 cc.; sp. gr. 1.17)	50	15	3.5	"	"	74	227
	H <sub>3</sub> PO <sub>4</sub> (2.0 cc.; sp. gr. 1.75)	70	25	3.5	"	"	90	230
	KBr, H <sub>2</sub> SO <sub>4</sub> (d)	55-60	25	3.5	"	"	91	230
			10	4.0	3,5-Dibromosalicylic acid	Glacial acetic acid	77	227-228
Glacial acetic acid (10.0 cc.)	Not added	70-75	90	3.5	3,5-Diiodosalicylic acid	Aqueous acetone	74	225
	H <sub>2</sub> SO <sub>4</sub> (1.0 cc.; sp. gr. 1.83)	70-75	120	3.5	(e)	"		
		(c)	24 hours	3.5	(e)	"		
	HNO <sub>3</sub> (1.0 c.c.; sp. gr. 1.42)	(c)	24 hours	3.5	(f)	"		
	HCl (2.0 cc.; sp. gr. 1.17)	60-65	15	2.3	3,5-Diiodosalicylic acid	Aqueous acetone	76	216
	HBr (5.0 cc.; sp. gr. 1.7)	70	75	4.0	3,5-Dibromosalicylic acid	Glacial acetic acid	66	222
						Glacial acetic acid	68	227

(a) For comparison the product was recrystallized once from the solvent listed.

(b) The melting point of 3,5-diiodosalicylic acid has been described, 220, 228-230, 235-236 °C.

(c) Room temperature.

(d) Iodine was not added to this reaction mixture.

(e) 3.0-3.5 g. of iodine unreacted.

(f) 2.8-3.0 g. of iodine unreacted.

that in addition to phenols, ethers, and hydrocarbons, ketones, tyrosine, nitrotyrosine, and methoxybenzaldehydes are readily iodinated under similar conditions (unpublished data).

### III. EXPERIMENTAL

#### (i) *3,5-Diiodosalicylic Acid*

To a vigorously agitated mixture of salicylic acid (4.0 g.), powdered iodine (7.4 g.), sulphuric acid (2.0 cc., sp. gr. 1.83), and 95% ethanol (20.0 cc.), 30% hydrogen peroxide (4.5 cc.) was added in 0.5 cc. portions during 10 minutes, the temperature being maintained at 60 °C. When 3.0 cc. of peroxide had been added a violent reaction occurred, the temperature rose to 85 °C., and all the iodine dissolved. When the reaction had subsided the addition of peroxide was completed. After shaking for a further 5 minutes at 60–65 °C., adding water (80 cc.), and cooling, the crystalline, pink precipitate was collected at the pump, washed with water, drained thoroughly, and washed with three 4.0 cc. portions of cold glacial acetic acid, most of the red colouring matter thereby being removed. After a final washing with water and draining, the off-white product was dissolved in boiling acetone (15.0 cc.). On adding water (80 cc.), slowly and with stirring, 3,5-diiodosalicylic acid separated as almost colourless needles, melting at 234 °C. (lit. 235–236 °C.). Yield 10.43 g.; 91%.

Found: I, 65.6%.

Calculated for  $C_7H_4O_3I_2$ : I, 65.1%.

#### (ii) *5-Iodosalicylic Acid*

Hydrogen peroxide (30%, 2.8 cc.) was added gradually during 5 minutes to an agitated suspension of powdered iodine (5.52 g.) and salicylic acid (6.0 g.) in glacial acetic acid (40.0 cc.), and concentrated hydrochloric acid (31%, 6.0 cc.), maintained at 60 °C., until the colour of iodine was discharged. The temperature was then raised to 75–80 °C. and agitation was continued for a further 5 minutes. After adding water (150 cc.) and cooling, the granular, off-white precipitate was collected, dissolved in boiling 5% sodium hydroxide solution (30.0 cc.), and added to a boiling solution of barium hydroxide octahydrate (7.0 g.) in water (200 cc.). Boiling was stopped immediately, the mixture stirred for 2 minutes and the white, precipitated barium salt filtered from the hot filtrate (A). The granular acid, which separated when a solution of the barium salt in boiling glacial acetic acid (20.0 cc.) was added to a cold 5% hydrochloric acid solution (100 cc.), was dissolved once again in hot 5% sodium hydroxide solution (15.0 cc.), the solution added to a boiling aqueous solution (100 cc.) of barium hydroxide octahydrate (4.0 g.), the precipitated barium salt separated from the hot filtrate (B) and dissolved in boiling glacial acetic acid (15.0 cc.). When this solution was added to a cold 5% hydrochloric acid solution (100 cc.), a pure white, granular product, melting at 188–190 °C., was obtained (3.75 g.). Recrystallization of this product from glacial acetic acid with the aid of a little charcoal gave colourless needles of 3,5-diiodosalicylic acid, melting at 230 °C. (1.82 g.). The combined filtrates (A) and (B) were acidified with concentrated hydrochloric acid, cooled, and filtered. A solution of the acid, thus obtained, in boiling 2N sodium carbonate solution (45 cc.) was treated with charcoal, filtered, and acidified with hydrochloric acid. The white, granular product was collected immediately, washed with hot water, and dried (6.10 g.; m.p. 170–172 °C.). Two recrystallizations from glacial acetic acid gave long, colourless needles of 5-iodosalicylic acid, melting at 198 °C. (lit. 198–199 °C.). Yield 2.78 g.; 24%.

Found: I, 48.5%.

Calculated for  $C_7H_5O_3I$ : I, 48.1%.

#### (iii) *3,5-Dibromosalicylic Acid*

(a) An agitated mixture of salicylic acid (2.0 g.), powdered potassium bromide (3.45 g.), sulphuric acid (2.0 cc., sp. gr. 1.83), alcohol (10.0 cc.), and water (5.0 cc.), maintained at 55–60 °C., was treated gradually with 30% hydrogen peroxide (4.0 cc.) during 5 minutes. Agitation at this temperature was continued for a further 5 minutes during which time a

crystalline precipitate separated. After adding excess water and cooling, the precipitated bromo compound (4.01 g.), melting at 212 °C., was collected and recrystallized from glacial acetic acid, 3,5-dibromosalicylic acid being obtained as colourless needles, melting at 227–228 °C. (lit. 228 °C.). Yield 3.30 g.; 77%.

Found: Br, 53.3%.

Calculated for  $C_7H_4O_3Br_2$ : Br, 54.1%.

(b) Hydrogen peroxide (30%, 4.0 cc.) was added in 0.5 cc. portions during 20 minutes to a suspension of powdered iodine (3.7 g.) in glacial acetic acid (10.0 cc.) and hydrobromic acid (5.0 cc., sp. gr. 1.7). The reaction was exothermic and after adding 2.0 cc. of peroxide all the iodine had dissolved. Salicylic acid (2.0 g.) was then added to the clear, red solution and the temperature raised to 70 °C. on the water-bath. Within a few minutes crystals of iodine began to separate from the reaction mixture. After 45 minutes at this temperature the mixture was cooled, diluted with water (15.0 cc.), and extracted with ether. After decolourizing the ether extract with aqueous sodium thiosulphate the solvent was removed. The light yellow, crystalline residue (4.15 g.), melting at 214 °C., was recrystallized from glacial acetic acid, 3,5-dibromosalicylic acid being obtained as colourless needles, melting at 227 °C. Yield 2.50 g.; 60%. Found: Br, 53.9%.

(c) A mixture of salicylic acid (6.0 g.), powdered iodine (11.1 g.), hydrobromic acid (14.0 cc., sp. gr. 1.7), and glacial acetic acid (30.0 cc.) was treated with 0.5 cc. portions of 30% hydrogen peroxide (12.0 cc.) during 30 minutes. When all the iodine had dissolved the red solution was heated to 70 °C. Crystals of iodine immediately separated. After 45 minutes the mixture was cooled and the dibromosalicylic acid, melting at 214–215 °C. (12.7 g.), isolated as in (b). Recrystallization of this product from glacial acetic acid gave colourless needles, melting at 227 °C. Yield 8.71 g.; 68%. Found: Br, 53.7%.

(iv) *2,6-Diiodo-4-Nitrophenol*

To a mixture of sulphuric acid (1.0 cc., sp. gr. 1.83), powdered iodine (3.66 g.), and *p*-nitrophenol (2.0 g.) in alcohol (10.0 cc.), 30% hydrogen peroxide (3.5 cc.) was added gradually during 10 minutes. A vigorous reaction, with some decomposition of the peroxide, occurred with each addition, so that external heating was unnecessary. After shaking for a further 5 minutes water (40 cc.) was added, the mixture cooled, and the almost pure diiodo compound, melting at 154 °C., collected (5.13 g.). One recrystallization from glacial acetic acid gave yellow needles of 2,6-diiodo-4-nitrophenol, melting at 156 °C. (lit. 155–156 °C.). Yield 4.60 g.; 82%.

Found: I, 65.3%.

Calculated for  $C_6H_3O_3NI_2$ : I, 65.0%.

(v) *α-Iodo-β-Ethoxy-β-Phenylpropionic Acid*

Hydrogen peroxide (30%, 4.30 cc.) was added during 15 minutes to an agitated mixture of cinnamic acid (4.0 g.), powdered iodine (3.44 g.), sulphuric acid (2.0 cc., sp. gr. 1.83), and 95% ethanol (20.0 cc.), maintained at 50–60 °C. After standing at room temperature for 40 hours most of the iodine had reacted. The red solid, which separated on adding water (80 cc.), was collected and the aqueous filtrate extracted with ether (A). The solid was suspended in warm N sodium carbonate solution (80 cc.), a small amount of insoluble, red oil being extracted with the ether solution (A). After expelling dissolved ether by warming, the sodium carbonate solution was acidified with hydrochloric acid. The light yellow precipitate, dried at 100 °C., melted at 137 °C., and consisted of the almost pure iodoethoxy acid. Yield 4.11 g.; 48%. Recrystallization of this product from water gave colourless needles of the monohydrate of *α*-iodo-*β*-ethoxy-*β*-phenylpropionic acid, melting at 141 °C. (lit. 137–138 °C.).

Found: I, 37.9%.

Calculated for  $C_{11}H_{13}O_3I.H_2O$ : I, 37.6%.

Using methanol as solvent the dihydrate of *α*-iodo-*β*-methoxy-*β*-phenylpropionic acid, melting at 169 °C. (lit. 168–169 °C.), was obtained by a similar process.

Found: I, 37.0%.

Calculated for  $C_{10}H_{11}O_3I.2H_2O$ : I, 37.1%.



(a) *The Iodination of Hydrocarbons*

A description of the general procedure is given in one instance, the other results are tabulated in Table 2.

TABLE 2

Hydro-carbon	Reaction Period (hr.)	Product	Yield (%)	Boiling Point (°C.)	Boiling Point (lit. °C.)	Iodine	
						Cal- culated (%)	Found (%)
<i>m</i> -Xylene ..	0.5	4-Iodo- <i>m</i> -xylene	74	228–231	232	54.7	54.5
<i>p</i> -Xylene ..	1.0	2-Iodo- <i>p</i> -xylene	64	229–230	229	54.7	55.1
Toluene ..	2.0	<i>o</i> - and <i>p</i> -Iodo-toluene	69	209–211	211	58.2	57.6
Benzene ..	5.0	Iodobenzene	12	180–190	186–188	62.3	61.2

(i) *The Iodination of Toluene*

A large flask containing a mixture of toluene (13.0 g.), powdered iodine (17.9 g.), sulphuric acid (6.0 cc., sp. gr. 1.83), and alcohol (50 cc.), was fitted with a reflux condenser and suspended over a hot water-bath so that the temperature of the contents was maintained at about 70 °C. Hydrogen peroxide (30%, 24.0 cc.) was added through the condenser in 15 portions during 1 hour, the mixture being vigorously agitated after each addition, until the colour of iodine was discharged. The brownish red reaction mixture was then refluxed gently on the water-bath for 1 hour, a further quantity of hydrogen peroxide (10.0 cc.) being added during this period. After cooling, diluting with water (200 cc.), and making alkaline with sodium hydroxide, the lower, yellow oily layer was separated and steam distilled from a slightly alkaline solution. The oil which separated in the distillate was collected, dried over anhydrous sodium sulphate, and distilled at atmospheric pressure, the iodotoluene fraction, boiling at 208–210 °C., being collected. Yield 21.1 g.; 69%.

Similar results were obtained when concentrated nitric acid was employed in the place of sulphuric acid.

## IV. ACKNOWLEDGMENTS

The author is indebted to Mr. G. K. Hughes, Department of Organic Chemistry, University of Sydney, for his kindness in criticizing the material presented in this paper, and to Dr. C. R. Barnicoat, Massey Agricultural College, New Zealand, for providing facilities to carry out this work.

## V. REFERENCES

- (1) WOOLLETT, G. H., and JOHNSON, W. W.—“Organic Syntheses.” Coll. Vol. 2, p. 343. (John Wiley and Sons: New York, 1944.)
- (2) BLOCK, P., and POWELL, G.—*J. Amer. Chem. Soc.* **64**: 1070 (1942).
- (3) LUCAS, H. J., and KENNEDY, E. R.—“Organic Syntheses.” Coll. Vol. 2, p. 351. (John Wiley and Sons: New York, 1944.)
- (4) DATTA, R. L., and CHATTERJEE, N. R.—*J. Amer. Chem. Soc.* **39**: 435 (1917).
- (5) ELBS, K., and JAROSLAWZEW, A.—*J. prakt. Chem.* **88**: 92 (1913).
- (6) KLAGES, A. ET AL.—*J. prakt. Chem.* **61**: 311 (1900); *J. prakt. Chem.* **65**: 564 (1902).
- (7) MARSH, J. E.—*J. Chem. Soc.* **1927**: 3164 (1927).
- (8) JURD, L.—*Aust. J. Sci. Res. A* **2**: 595 (1949).
- (9) LEULIER, A., and PINET, L.—*Bull. Soc. Chim.* **41**: 1362 (1927).
- (10) JACKSON, E. L., and PASIUT, L.—*J. Amer. Chem. Soc.* **50**: 2249 (1928).

# THE ULTRAVIOLET ABSORPTION SPECTRA OF THE ACRIDONE ALKALOIDS

## I. COMPOUNDS CONTAINING THE ACRIDONE NUCLEUS

By R. D. BROWN\* and F. N. LAHEY†

[*Manuscript received December 22, 1949*]

### *Summary*

The ultraviolet spectra of the acridone alkaloids and some derivatives have been studied. The various bands in the spectra of these compounds are interpreted by starting with theoretical treatments of the spectrum of anthracene and tracing spectral changes through acridine and phenazine to acridone. An interesting relation between the  $\pi$ -electron density of a given position in the acridone ring and the shift in wavelength of absorption bands due to an alkoxyl derivative is noted and the spectra of the alkaloids are shown to comply with this relationship. Compounds containing a 4-hydroxy-10-methylacridone structure are found to have unusual spectra and a possible explanation of this is proposed.

## I. INTRODUCTION

The chemistry of the alkaloids acronycine, melicopine, melicopidine, melicopicine, evoxanthine, and their derivatives is known in some detail(1-3) and the structures of most of these have been established. The ultraviolet absorption spectra of these alkaloids and derivatives were studied in conjunction with the chemical investigations, and in general it may be said that they confirm the chemical evidence. However, the spectra are of interest in another direction ; it has proved possible to interpret the various band systems constituting the spectra using for the most part theoretical concepts based upon the molecular-orbital theory of quantum mechanics. As this method of interpretation does not seem to have been attempted before in the case of such relatively complicated molecules, the spectra will be discussed in some detail. The present paper contains the experimental results and a qualitative discussion ; a more detailed account of the relevant theoretical considerations will be published by one of us (R.D.B.) shortly.

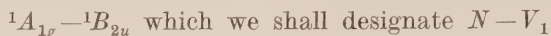
## II. GENERAL INTERPRETATION OF RESULTS

We may start the discussion of the spectra of compounds containing the acridone nucleus by considering the spectrum of anthracene ; it consists(4, 5) of two distinct band systems, the one of longer wavelength having a nicely resolved vibrational structure. If we employ the coordinate system shown in

\* Organic Chemistry Laboratory, University of Melbourne.

† Present address : Department of Chemistry, University of Queensland.

Figure 1, then it may be shown by a molecular-orbital treatment(5) that the first two permitted electronic transitions are



both being symmetry-permitted transitions.\* The calculations also predict a forbidden transition to an upper  ${}^1B_{3g}$  level lying roughly in the same spectral region as the second permitted transition. We shall refer to this transition as the  $N-V_{2f}$  transition. That the molecular-orbital treatment, contrary to the Lewis-Calvin theory of colour, correctly predicts the nature of the first two electronic transitions in the anthracene spectrum is demonstrated by other experimental evidence (6, p. 263); clearly the very feeble absorption of the  $N-V_{2f}$  transition is masked by the strong  $N-V_{2p}$  transition in the same region.

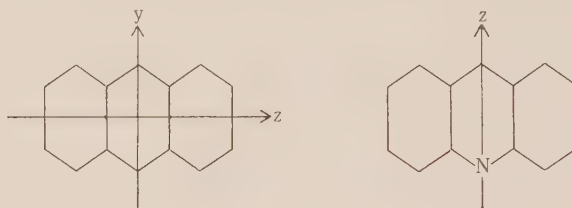


Fig. 1

A comparison of the anthracene spectrum with that of acridine and phenazine, shown in Figure 2, is rather interesting.† The anthracene spectrum is practically unaltered by the successive introduction of two nitrogen atoms at the meso positions. This is rather unexpected from interpretations of the spectrum of anthracene based upon the alternative quantum-mechanical method—the HLSP method—as given by Jones(7) and by Coulson(6), because the introduction of nitrogen atoms would be expected to stabilize structures such as :



which contribute largely to the  ${}^1B_{2u}$  state involved in the longest wavelength absorption band and so should shift the position of the band to longer wavelengths. One probable explanation is that these ionic structures contribute in a similar degree to the  ${}^1A_{1g}$  ground state, and thus have a very similar effect upon the energies of the two levels. This explanation is borne out in two ways : (i) first-order perturbation theory (R. D. Brown, unpublished data) for an

\* For an account of electronic selection rules see Sponer, H., and Teller, E.—*Rev. Mod. Phys.* **13**: 75 (1941).

† The spectra shown in Figure 2 were derived from the results of Radulescu and Ostrogovich(5).

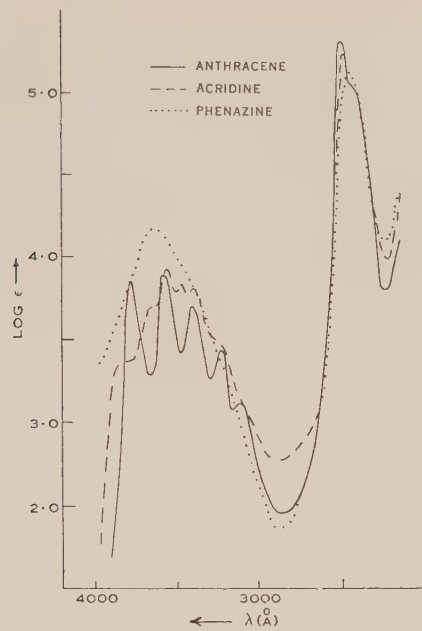


Fig. 2

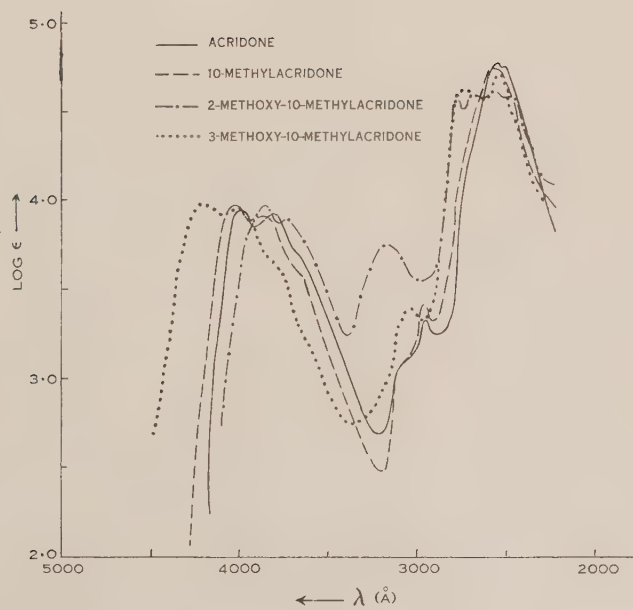


Fig. 3



alternant hydrocarbon will always predict no change in the transition energy, and it is well known (see, for example, 8) that the effect of introduction of a heteroatom may be allowed for approximately by a perturbation calculation; (ii) calculations of  $\pi$ -electron densities and bond orders for acridine(9) indicate that the above ionic structures contribute appreciably to the ground state.

One slight difference which can be noticed is that the longer wavelength band, which we shall henceforth term the  $V_1$  band, is of slightly greater intensity, and the shorter wavelength band, which will now be referred to as the  $V_{2p}$  band, is of slightly less intensity in the spectrum of acridine than in anthracene; this effect is more apparent in the spectrum of phenazine. Theoretical computations show(9) that the  $\pi$ -electrons tend to accumulate around the heteroatoms so this may increase the transition moment for the  $y$ -direction and decrease it for the  $z$ -direction.

It may be noticed that the  $V_{2f}$  bands in acridine and phenazine are still masked by the more intense  $V_{2p}$  bands. In acridine and related molecules of symmetry  $C_{2v}$  the  $V_{2f}$  band is due to a permitted transition, i.e. the forbidden transition of anthracene has become permitted due to the perturbation by the nitrogen atom. Such *perturbation-permitted* bands will be of intensity intermediate between that of a symmetry-permitted band and a symmetry-forbidden band. The perturbation-permitted  $V_{2f}$  band will be referred to again below.

Upon introduction of a hydroxyl group to give 5-hydroxyacridine (acridone) the  $V_1$  band is shifted to longer wavelengths, and the  $V_{2p}$  band is only slightly shifted in position. At the same time the  $V_1$  band is slightly increased in intensity and the  $V_{2p}$  band suffers a large decrease in intensity (Fig. 3). The intensity change is analogous to that observed in going over to phenazine (see above), and, as theoretical calculations show (R. D. Brown, unpublished data) again that the  $\pi$ -electrons tend to congregate around the nitrogen and hydroxyl group at the expense of the rest of the molecule, this is not unexpected.

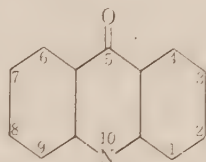
The shifts in positions of the bands upon introduction of hydroxyl is due to the stabilization of the ionic forms previously considered, particularly those which would contribute to a transition moment in the  $y$ -direction and thus affect the  $V_1$  band rather than the  $V_{2p}$  band. (The  $V_1$  band has a transition moment in the  $y$ -direction, the  $V_{2p}$  band in the  $z$ -direction in anthracene(7).) This explanation presupposes that the ionic forms do not contribute so strongly to the ground state of acridone (contrast the case of acridine considered above); this is confirmed by the low bond order of the carbon-oxygen bond in acridone inferred both from theoretical calculations (R. D. Brown, unpublished data) and experimentally by a study of the ultraviolet absorption spectra of a series of hydroxyacridines(10).

The absorption spectrum of acridone exhibits a new band in the region of 3000 Å which is not observed in the spectra of anthracene or acridine. Many carbonyl compounds exhibit absorption in this spectral region so the most tempting interpretation is perhaps an  $^1A_1 \rightarrow ^1A_2$  transition, i.e. a symmetry-forbidden transition. However, this possibility is ruled out by the intensity of the absorption, which is far too great for a transition of this nature. Consequently it seems much more likely that this band is due to a perturbation

permitted transition, as mentioned above. When the symmetry of a molecule is perturbed from  $V_h$  to  $C_{2v}$  the levels of symmetries  ${}^1B_{2u}$ ,  ${}^1B_{1u}$ ,  ${}^1B_{3g}$  in the former become (after making due allowance for change in coordinates as in Fig. 1)  ${}^1A_1$ ,  ${}^1B_2$ ,  ${}^1B_2$ , respectively, in the latter. The corresponding transition moments are in the latter cases in the directions  $z$ , approximately  $y$ , approximately  $y$ , respectively. We shall interpret the band around 3000 Å in the spectrum of acridone as due to the transition  ${}^1A_1 \rightarrow {}^1B_2$ , and for correlation with the other spectra discussed above this will henceforth be referred to as  $N-V_{2f}$ .

The spectrum of 10-methylacridone (Fig. 3) is very similar to that of acridone, the bands being shifted slightly to longer wavelengths in the spectrum of the former compound. This red shift accompanying N-methylation is also observed in the case of 4-quinolone and some of its derivatives and is due to the effect of hyperconjugation(11) upon the energies of the electronic levels involved.

The  $V_{2f}$  band becomes more prominent in the spectra of the methoxyacridones (Figs. 3 and 4). The compounds studied were 10-methyl-2-methoxy-, 10-methyl-3-methoxy-, and 10-methyl-4-methoxyacridone (the numbering system employed is given in I). Each of these methoxyls reduces the symmetry of 10-methylacridone from  $C_{2v}$  to approximately  $C_s$ , and it is this increased perturbation to the electronic symmetry which makes the  $V_{2f}$  band still less "forbidden", i.e. considerably increases its intensity.



(I)

The introduction of methoxyl also affects the positions of the other bands, particularly the position of the  $V_1$  band; a methoxyl group in the 3 position shifts the  $V_1$  band to longer wavelengths, while in both the 2 and the 4 positions the shift is to shorter wavelengths.

Comparison of the absorption spectra of 4-hydroxy-10-methylacridone with that of the 4-methoxy compound (Fig. 4) reveals that the main difference is in the position of the  $V_{2f}$  band. This is further evidence of the close connection of this  $V_{2f}$  band with the presence of oxy groups in the molecule.

It will be noticed that the  $V_{2p}$  band shows increased vibrational structure in these methoxy and hydroxy derivatives. In the 4-hydroxy derivative in particular a frequency of about 1300  $\text{cm}^{-1}$  is noticeable and such a frequency is likely to be associated with the vibration of an OH group(12-14). This interpretation is supported by the fact that the group is on one side of the molecule and thus would be more likely to vibrate in a transition with a moment in the  $z$ -direction. However, such an assignment is at the best very tentative.

The spectrum of 2,4-dihydroxy-10-methylacridone is shown in Figure 4. The  $V_1$  band is in the same position as that of 10-methylacridone, and of slightly

lower intensity, the  $V_{2p}$  band being shifted in the expected direction. Now in the case of the 4-hydroxy derivative the  $V_1$  band was in the same position as in the corresponding methoxy compound so it would appear that the  $V_1$  band of the dimethoxy compound lies in the same spectral region as that of 10-methylacridone. On the scheme given above we should expect 2,4-dimethoxy substituents to shift the  $V_1$  band to shorter wavelengths, together with a shift of the  $V_{2p}$  band to slightly longer wavelengths, this latter being the general direction of shift of the  $V_{2p}$  band in all of the methoxy derivatives so far considered. Consequently it would seem that the  $V_1$  band of 2,4-dimethoxy-10-

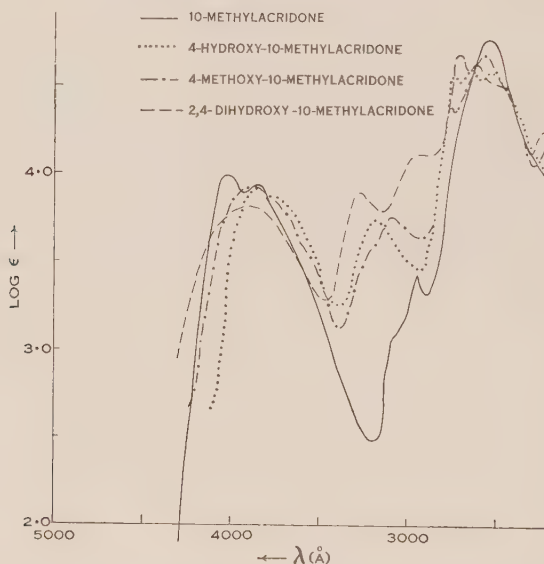
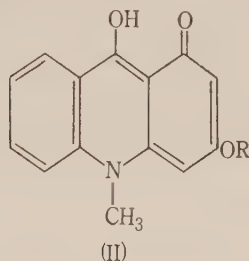


Fig. 4

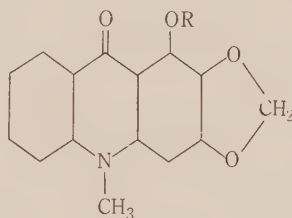
methylacridone is not in the anticipated position. The absorption spectrum of this compound has not been determined, but the qualitative observation that the dihydroxy compound is yellow while the dimethoxy compound is colourless indicates that the  $V_1$  band of the latter actually lies at shorter wavelengths, and thus at shorter wavelengths than the  $V_1$  band of 10-methylacridone as was predicted.

The phenomenon of a 4-hydroxy-10-methylacridone having its  $V_1$  band at longer wavelengths than the corresponding methoxy compound when there is an oxy-substituent in the 2 position will be met with again in several other acridone derivatives. For reasons which will appear then it will henceforth be termed the *nor*-effect. It may possibly be connected with the possibility of the system assuming the tautomeric form (II). Presumably such a tautomer is only stable when a 2-oxy-substituent is present. It may be noted that the parent compound, 10-methyl-4-acridone, is blue(15).



Two other features are present in the spectrum of 2,4-dihydroxy-10-methylacridone. Firstly the  $V_{2f}$  band which has progressively increased in intensity from acridone to the mono-substituted-10-methylacridones, has now been replaced by two bands in the same spectral region with intensity above that of the  $V_1$  band in the dihydroxy derivative. This, of course, fits in well with the interpretation of the  $V_{2f}$  band given previously.

The second feature is again the presence of vibrational structure in the  $V_{2p}$  band. It is interesting to note that the two submaxima have a separation of  $1200\text{ cm.}^{-1}$ , in satisfactory agreement with that found for the 4-hydroxy compounds.



(III)  $R = \text{CH}_3$

(IV)  $R = \text{H}$ .

The electronic spectra of the alkaloid evoxanthine (III) and of *nor*-evoxanthine (IV) are shown in Figure 5. The  $V_1$  band of evoxanthine almost coincides with that of 10-methylacridone; the  $V_{2p}$  band shows the characteristic shift to longer wavelengths.

In the spectrum of *norev*oxanthine the  $V_1$  band has undergone a considerable shift to longer wavelengths. This is a case of the *nor*-effect mentioned above; it will be noticed that there is an ether substituent present in the 2 position.\* Another feature of the spectrum is the slight shift of the  $V_{2f}$  band to longer wavelengths relative to evoxanthine. This has already been observed in the comparison of the spectra of 4-hydroxy- and 4-methoxy-10-methylacridone and is possibly connected with a lowering of the energy of the upper perturbed  ${}^1B_{3g}$  type level due to increased contributions from certain ionic structures involving the free hydroxyl group.

\* All the compounds so far known which exhibit the *nor*-effect contain an ether substituent in the 2 position. However, it is quite possible, of course, that an ether substituent in another position, or another type of substituent, would occasion the same phenomenon.



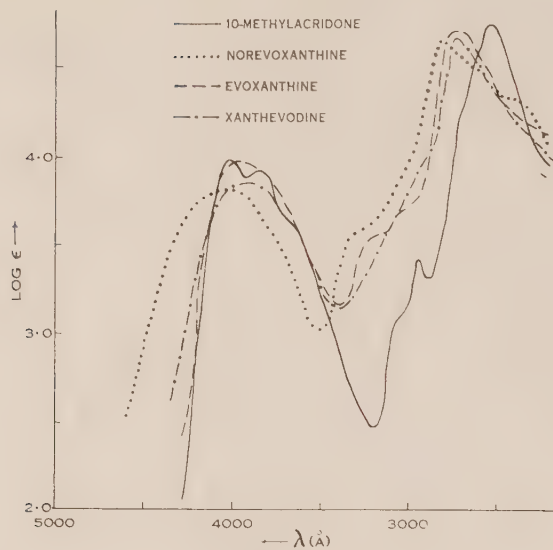


Fig. 5

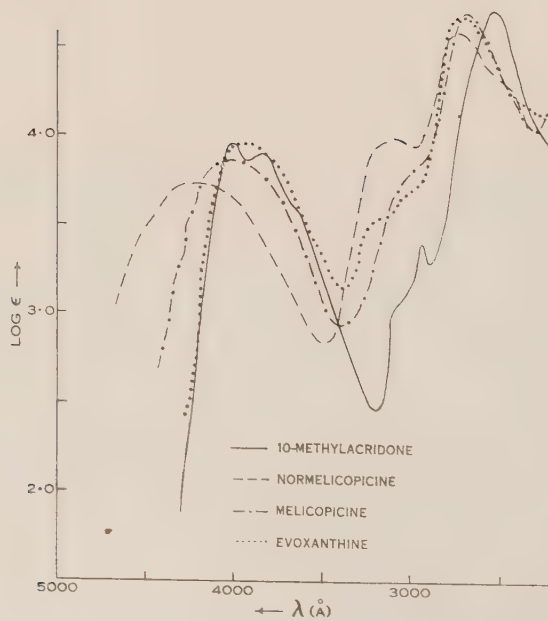
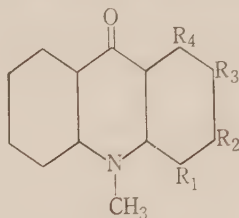


Fig. 6

The alkaloid melicopine is a 10-methylacridone with four methoxy groups, these being 1,2,3,4 (V). Such a close arrangement of methoxyl groups must occasion very great steric interactions, as may be seen from Fischer-Hirschfelder type models for example. Now for strong delocalization of the oxygen  $2p$  lone pair electrons into the acridone ring the carbon atom of the methyl group attached to the oxygen must lie close to the plane of the ring system. The strong steric interactions thus will sensibly prevent this conjugation.

In the melicopine (Fig. 6) spectrum both  $V_1$  and  $V_{2p}$  bands are shifted to longer wavelengths while the  $V_{2f}$  band is little affected. The lack of change in the latter band indicates that delocalization of the oxygen  $2p$  electrons is required for shift in its position (in all of the cases previously considered this delocalization was not prohibited by large steric effects). It will be noticed that the  $V_{2f}$  bands are at longer wavelengths in evoxanthine, again indicating the necessity for delocalization of the oxygen electrons.

The spectrum (Fig. 6) of *normelicopine* (VI), in addition to exhibiting the *nor*-effect in the  $V_1$  band, shows a distinct shift of the  $V_{2f}$  band to longer wavelengths. This is in accord with the above interpretation because demethylation of the 4-methoxyl group allows delocalization of the oxygen electrons due to the very much smaller steric effect of hydrogen compared with methyl.



- (V)  $R_1=R_2=R_3=R_4=\text{OCH}_3$   
 (VI)  $R_1=R_2=R_3=\text{OCH}_3$ ;  $R_4=\text{OH}$   
 (XII)  $R_1=R_3=\text{O.CO.CH}_3$ ;  $R_2=\text{OCH}_3$   
            $R_4=\text{OH}$   
 (XVI)  $R_1=R_3=\text{H}$ ;  $R_2=\text{O.C(CH}_3)_2\text{.CO}_2\text{H}$   
            $R_4=\text{OCH}_3$   
 (XXII)  $R_1=\text{H}$ ;  $R_3=\text{CH:CH.C(CH}_3)_2\text{.CH}_2$   
            $R_2=\text{OH}$ ;  $R_4=\text{OCH}_3$ .

The spectra of the alkaloids melicopidine (VIII), differing from evoxanthine only by the presence of an additional methoxyl group at position 1, and of melicopine (VII), isomeric with melicopidine, are shown in Figure 7. The most interesting feature of their spectra is the positions of the  $V_1$  bands. It seems possible to fit the observed shift of the  $V_1$  band of 10-methylacridone due to one or more oxy-substituents, as evinced by these and the preceding spectra into a single unified scheme. It is first necessary to consider more closely the shifts due to a single methoxyl substituent (Table 1). For these compounds the oxygen  $2p\pi$  electrons can delocalize because the methyl group is not prevented from lying coplanar with the acridone ring. Now comparison of the  $V_1$  bands of evoxanthine and melicopidine shows that a sterically hindered methoxyl in position 1 has little effect. This indicates that the underlying cause of the shifts

listed in Table 1 is the delocalization of the oxygen  $2p\pi$  electrons.\* If we accept this, then comparison of the  $V_1$  bands of melicopidine and melicopicine indicates that the shift to shorter wavelengths caused by methylenedioxy in position 2 is greater than the shift to longer wavelengths produced by the same substituent in position 3 (by 300  $\text{cm}^{-1}$ ). The fact that this difference is larger than that for an unhindered methoxyl (Table 1) points to the following order of effectiveness of substituents:

methylenedioxy    unhindered methoxy    hindered methoxy.

This is also the order of ease of delocalization of the oxygen  $2p\pi$  electrons. It will be greatest for methylenedioxy because the oxygen is held more rigidly in the position of greatest conjugation, and also the angle between the two  $2p\sigma$  oxygen bonds will be closer to  $\pi/2$  so that the  $2p\pi$  electrons will have less  $2s$  character than the corresponding electrons in methoxyl and hence will delocalize more readily. This is further evidence that the shifts are caused by delocalization rather than some other effect, e.g. the enhancement of the Coulomb integral of the carbon atom to which the oxy-substituent is attached (the so-called "inductive effect").

TABLE 1  
QUANTITATIVE DATA FOR SHIFTS DUE TO UNHINDERED METHOXYL

Position of Substituent	Band Shift ( $\text{cm}^{-1}$ )	Direction
1	—	—
2	1100	Shorter $\lambda$
3	1000	Longer $\lambda$
4	500	Shorter $\lambda$

Finally, the relative positions of the  $V_1$  bands of melicopine and melicopidine show that the effect of methylenedioxy in position 1 is to cause a greater shift (by 600  $\text{cm}^{-1}$ ) to longer wavelengths than the same substituent in position 3. Consequently the order of effectiveness of the various positions substituted in acridone must be

1 (+),    2 (—),    3 (+),    4 (—),

the positions 2 and 3 being very similar in magnitude, however.†

From comparison with the calculated  $\pi$ -electron densities for acridone (R. D. Brown, unpublished data) these shifts may be generalized thus: an oxy-substituent in a position of (locally) high  $\pi$ -density shifts the  $V_1$  band to longer wavelengths, in a position of (locally) low  $\pi$ -density, shifts this band to

\* It is unfortunate that 1-methoxy-10-methylacridone was not available for a direct comparison with the effect of an unhindered 1-methoxyl substituent. However, the self-consistent scheme presented here requires a substituent in this position to cause a greater shift than in any other position, and no other assumption leads to an alternative self-consistent picture.

† It should be remembered that the  $V_1$  band envelopes of the compounds discussed above have rather varying shapes so that the shifts of the band origins, which would form a better basis for discussion, are somewhat uncertain. Thus shifts of 100 to 200  $\text{cm}^{-1}$  may not be real, and the shifts referred to above should be regarded as uncertain by about this amount.

shorter wavelengths. The basic theoretical significance of this interesting generalization is not at present apparent. It is significant that such a correlation between shifts due to a substituent and  $\pi$ -electron densities also applies to quite different molecules, e.g. Plattner(16, 17) has found that alkyl substituents in the 1, 5, and 7 positions shift the absorption bands of azulene to longer wavelengths while in the 2, 3, 6, and 8 positions shifts occur in the opposite direction.

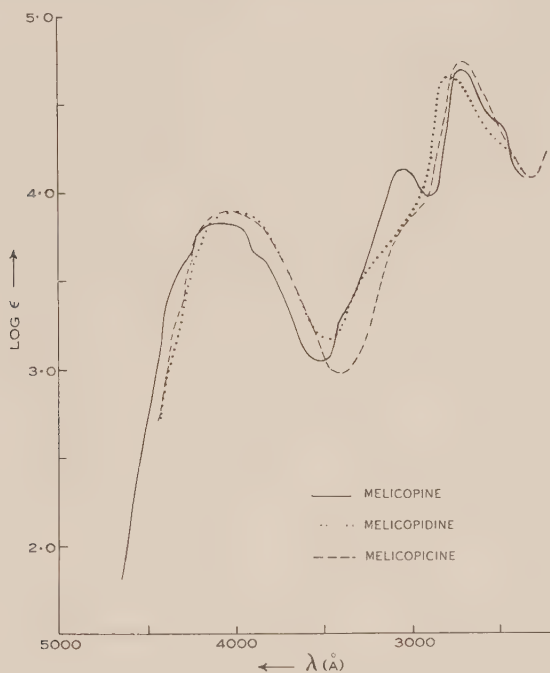


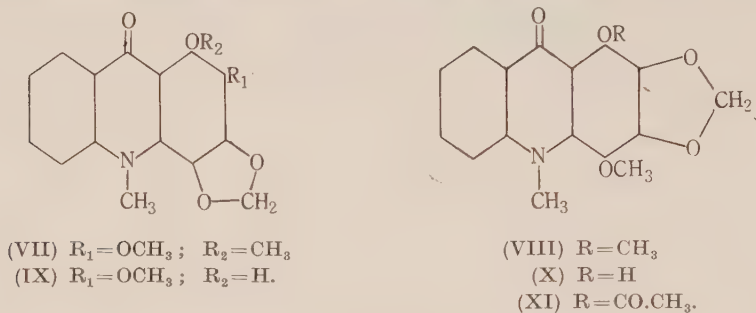
Fig. 7

These two sets of positions are those of (local) high and low  $\pi$ -densities respectively(18). It is interesting to observe in this connection that alternant hydrocarbons do not show such differential shifts, and it is just for these molecules that the LCAO molecular orbital treatment leads to a uniform distribution of  $\pi$ -electrons, non-alternants, and heterocyclic molecules having a non-uniform distribution.

It will be noted that when several oxy-substituents are present in the same molecule the effects mentioned above are approximately additive. For example, the  $V_1$  bands of melicopicine and melicopidine should lie in the same position as the  $V_1$  band of 10-methylacridone. In fact they lie at slightly longer wavelengths, the differences in position being scarcely significant. The  $V_1$  band of evoxanthine should likewise be close to that of 10-methylacridone; it appears to lie about  $100\text{ cm.}^{-1}$  to longer wavelengths whereas the expected shift is a similar amount in the opposite direction. Again this difference is not significant.



The spectra of the *nor*compounds of melicopine (IX), melicopidine (X), and melicopicine (VI) are shown in Figure 8. The relative positions of the three  $V_1$  bands are in excellent agreement with the above ideas. The  $V_{2f}$  bands in



each case are shifted to the red relative to the parent alkaloid, again indicating that delocalization of the oxygen electrons is responsible for this shift in this

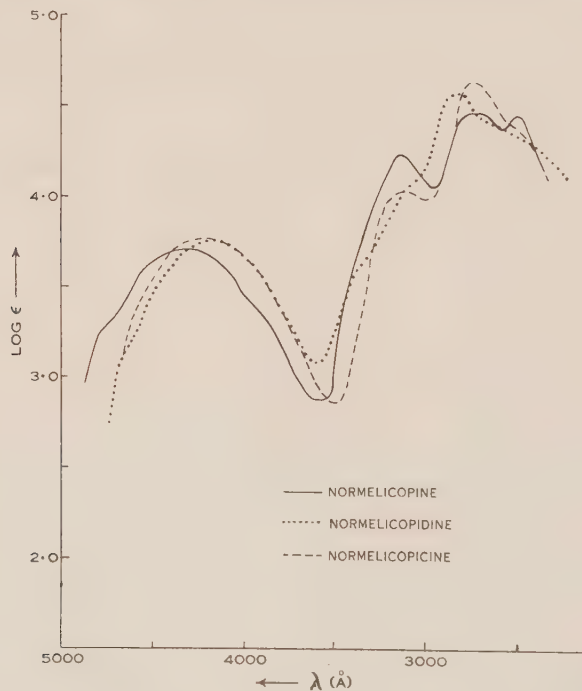


Fig. 8

band. It will be noticed that the  $V_{2f}$  bands of melicopine and melicopidine are at longer wavelengths than that of melicopicine due to delocalization of the methylenedioxy electrons and that in each case demethylation of the 4-methoxy substituent results in a further shift of these bands, to longer wavelengths.

It is clear from the above interpretations that the differences in the spectrum of a hydroxy compound from that of the corresponding methoxy compound are

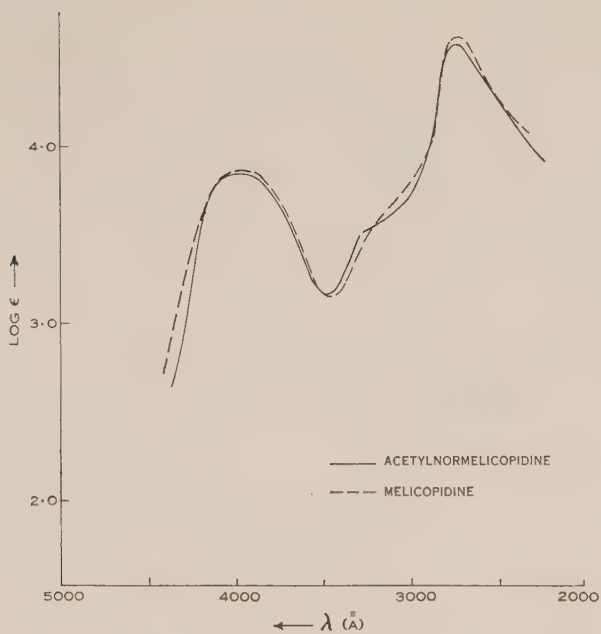


Fig. 9

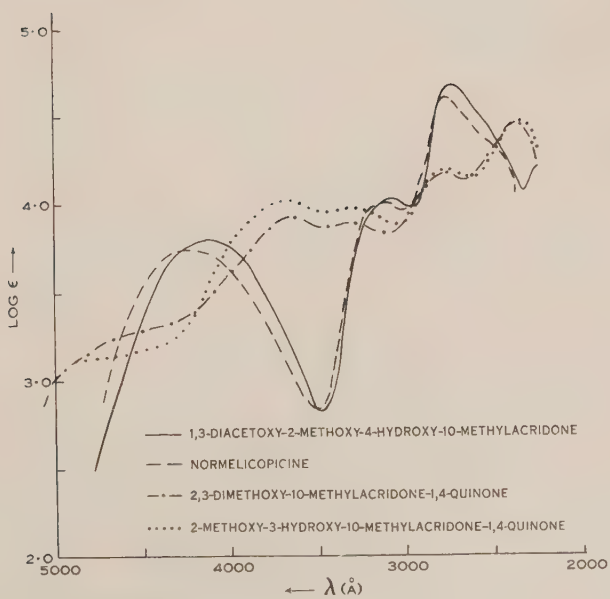
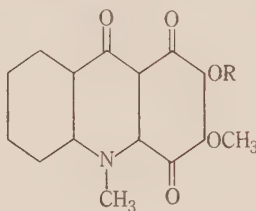


Fig. 10

due to two major factors: (i) lack of steric prevention of delocalization of the  $2p$  oxygen electrons as reflected mainly in the  $V_{2f}$  band; (ii) the *nor*-effect due to the assumption of a 4-acridone type structure. Consequently we should expect the replacement of methoxy by acetoxy to have but little effect upon the absorption curve. One example of this is the spectrum of acetyl*normelicopidine* (XI) which is compared with the spectrum of *melicopidine* in Figure 9. The spectra are seen to be almost identical.

A further example of this equivalence is provided by the spectrum of 1,3-diacetoxy-2-methoxy-4-hydroxy-10-methylacridone (XII). This would be expected to have a similar spectrum to *normelicopidine*, and it is seen (Fig. 10) that this is so.

The last two derivatives of the *Melicope* alkaloids to be considered are the compounds 2,3 - dimethoxy - 10 - methylacridone - 1,4 - quinone (XIII) and 2-methoxy-3-hydroxy-10-methylacridone-1:4-quinone (XIV). The ultraviolet

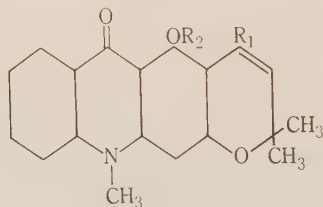


(XIII)  $R = \text{CH}_3$

(XIV)  $R = \text{H}$ .

absorption spectra of these compounds are shown in Figure 10. The absorption curves are quite different from the type previously considered, the number of band systems present being difficult to determine. The increased complexity is of course due to the introduction of a quinonoid structure and the interpretation of the spectra is difficult when the spectrum of the parent quinone is not available. However, the lack of any appreciable change accompanying demethylation of the 3-methoxyl indicates that the major spectral effect of this group is inductive.

The spectrum of another alkaloid, acronycine (XV), is shown in Figure 11. This has been shown (1) to be a 4-methoxy-10-methylacridone with a dimethyl pyran ring fused to the acridone system so that the oxygen of the former ring



(XV)  $R_1 = \text{H}$ ;  $R_2 = \text{CH}_3$

(XVIII)  $R_1 = \text{H}$ ;  $R_2 = \text{H}$

(XX)  $R_1 = \text{Br}$ ;  $R_2 = \text{CH}_3$

(XXI)  $R_1 = \text{Br}$ ;  $R_2 = \text{H}$ .

is attached to the 2 position of the latter. Thus the spectrum should be compared with that of 2,4-dimethoxy-10-methylacridone. The spectrum of monobasic acronycinic acid (XVI) (Fig. 14) should be very similar to this compound and a comparison with acronycine shows that the presence of a pyran double bond conjugated with the acridone system in the 3 position\* causes a small shift of the  $V_1$  band to longer wavelengths with accompanying shifts in the  $V_{2f}$  and  $V_{2p}$  bands, the most marked shift being in the last named band. It is also of interest that the  $V_1$  bands of acronycine and of 2,4-dihydroxy-10-methylacridone lie

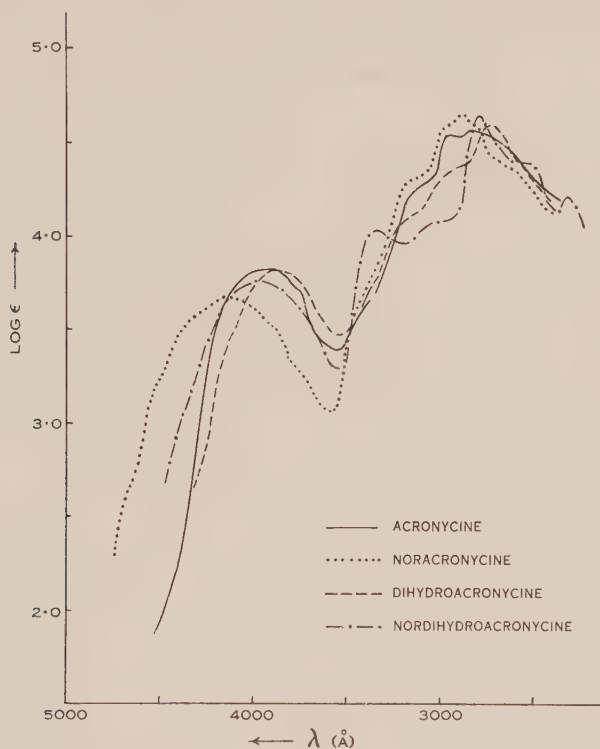


Fig. 11

at the same position showing that the shift due to the *nor*-effect (about 1100  $\text{cm}^{-1}$  on the average) is similar to that due to conjugation of the pyran double bond with the acridone system.

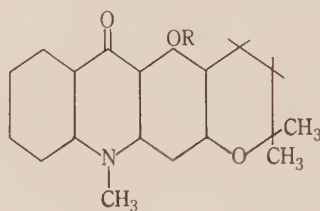
The spectrum (Fig. 11) of dihydroacronycine (XVII) is considerably closer to that of acronycinic acid, the only noticeable difference being that the  $V_1$  and  $V_{2p}$  bands are shifted slightly to longer wavelengths in the former, illustrating the much smaller effect of hyperconjugation upon these bands.

The spectrum (Fig. 11) of *nor*acronycine (XVIII) again exhibits the *nor*-effect. A comparison with the electronic spectrum of 2,4-dihydroxy-10-methylacridone shows that the  $V_1$  and  $V_{2p}$  bands lie at considerably longer wavelengths

\* It has not yet been definitely established whether the attachment is to the 1 or 3 position. Unpublished results favour position 3 as the point of attachment.



in the *nor*acronycine spectrum again showing the marked effect of conjugation of the chromene double bond. The *nor*-effect is also exhibited by *nordi*hydroacronycine (XIX) and the spectrum of the latter (Fig. 11) shows a very great similarity to that of 2,4-dihydroxy-10-methylacridone, the only noticeable differences, apart from intensity differences, being that the  $V_1$  and  $V_{2b}$  bands lie at slightly longer wavelengths in the former spectrum due to hyperconjugation of the dihydropyran methylene. Again the difference in position is greater for the  $V_{2b}$  band ( $700\text{ cm.}^{-1}$  compared with  $100\text{ cm.}^{-1}$ ).



(XIX)  $R=H$

(XVII)  $R=CH_3$ .

The spectra of bromacronycine (XX) and brom*nor*acronycine (XXI) are shown in Figure 12. The bromine atom in both cases is in the pyran ring and it is significant that the  $V_1$  band (transition moment in the  $y$ -direction) is almost

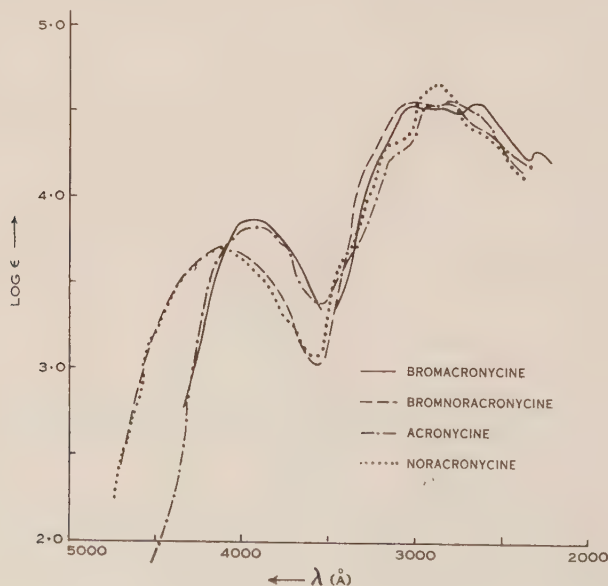


Fig. 12

identical with the  $V_1$  band of the unbrominated compound in both cases while the  $V_{2b}$  bands have been shifted to longer wavelengths and merged with the  $V_{2f}$  bands. From this result we can conclude that substituents at the side of the acridone system cause the greatest shift in position of bands with transition

moments across the longer dimension of the ring system and have less effect upon the bands with moments across the shorter direction of the ring. This is also borne out by the observations of the effect of the pyran double bond upon the spectrum of acronycine.

TABLE 2  
MAGNITUDES OF NOR-EFFECTS

Compound	(Cm. <sup>-1</sup> )
Melicopine .. .. .	1250
Melicopidine .. .. .	1100
Melicopicine .. .. .	1000
Acronycine .. .. .	1300
Bromacronycine .. .. .	1050
Dihydroacronycine .. .. .	700

The spectrum of bromnoraacronycine again exhibits the *nor*-effect. The magnitudes of the various *nor*-effects observed in the present study are tabulated in Table 2. It is rather surprising that the shift in position of the maximum

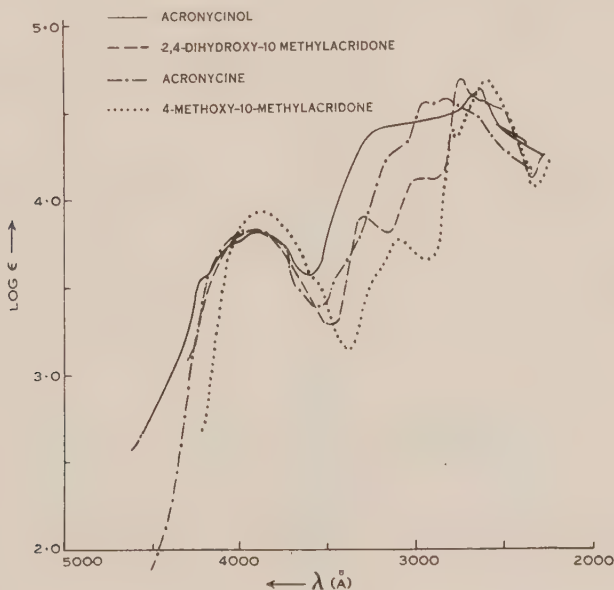


Fig. 13

(which is the quantity listed in Table 2) is so constant, particularly when it is remembered that the position of the maximum is by no means separated from the origin of the electronic band system by a constant amount.

The absorption spectrum of another derivative of acronycine—acronycinol (XXII)—is shown in Figure 13. One possible structure for this compound is a

2-hydroxy-4-methoxy-10-methylacridone with a 3-methyl-butadienyl side chain in the 3 position.\* The best test for such a structure would be comparison with the spectrum of 2-hydroxy-4-methoxy-10-methylacridone but the spectrum of this compound has not been determined. In view of the close similarity of

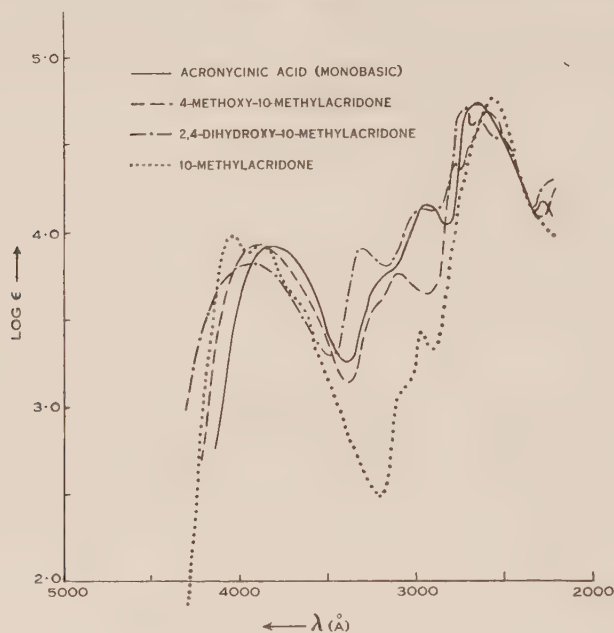


Fig. 14

the spectra of *nordihydroacronycine* and 2,4-dihydroxy-10-methylacridone we are justified in using the spectrum of the latter compound for comparison however. This last spectrum is very similar to that of acronycinol, especially in the  $V_1$  band, and the main change in the spectrum is the very marked shift of the  $V_{2p}$  band to longer wavelengths, merging with the lower intensity  $V_{2f}$

TABLE 3

## COORDINATES OF MAXIMA

Data applying to points of inflection are given in parentheses and indefinite inflections are denoted by ? or ??. The wavelengths are in Ångström units ( $10^{-8}$  cm.), the  $\log \epsilon$  apply to cell lengths in cm. and concentrations in mol. litre $^{-1}$ , and the frequencies are in wave numbers (cm. $^{-1}$ ).

Compound	Wavelength	Log $\epsilon$	Frequency
Acridone (I)	3995	3.95	25050
	3805	3.93	26300
	3080	3.05	32500
	2950	3.33	33900
	2545	4.78	39350
	2510	4.76	39900

\* See footnote p. 607.

TABLE 3 (Continued)

Compound	Wavelength	Log $\epsilon$	Frequency
10-Methylacridone	4040	3.98	24750
	3870	3.92	25900
	(3060) ?	(3.09) ?	(32700) ?
	2950	3.43	33900
	2560	4.76	39100
2-Methoxy-10-methylacridone	3865	3.97	25900
	3730	3.88	26900
	3180	3.75	31500
	2775	4.59	36000
	2690	4.63	37200
	2570	4.63	38900
3-Methoxy-10-methylacridone	4210	3.99	23800
	4020	3.95	24900
	3040	3.40	32900
	2730	4.63	36700
	2540	4.71	39400
4-Methoxy-10-methylacridone	3875	3.93	25800
	3100	3.76	32300
	2770	4.39	36100
	2585	4.68	38750
4-Hydroxy-10-methylacridone	3870	3.93	25850
	(3740)	(3.84)	(26800)
	3190	3.75	31400
	2770	4.56	36150
	2670	4.60	37500
	2575	4.58	38850
Xanthevodine	3930	2.41*	25500
	2740	3.22*	36550
Evoxanthine (III)	3970	3.97	25200
	(3210)	(3.54)	(31200)
	2745	4.71	36500
norEvoxanthine (IV)	4045	3.82	24700
	(3300)	(3.57)	(30350)
	2810	4.65	35600
	(2420)	(4.33)	(41300)
Acronycine (XV)	3920	3.83	25500
	(3760)	(3.72)	(26600)
	(3090)	(4.27)	(32400)
	2930	4.55	34150
	2810	4.57	35600

\* To obtain log  $\epsilon$  from these figures add log M/10.



TABLE 3 (Continued)

Compound	Wavelength	Log $\epsilon$	Frequency
<i>nor</i> Acronycine (XVIII) .. .. .	4130 (3130) (2945) 2885 2850	3.69 (4.31) (4.59) 4.65 4.66	24200 (32000) (34000) 34700 35100
Acronycinol (XXII) .. .. .	(4200) 3880 (3200) 2660 2580	(3.54) 3.82 (4.39) 4.58 4.55	(23800) 25800 (31300) 37650 38800
Melicopine (VII) .. .. .	4085 (3880) 3030 2705 (2520)	3.82 (3.64) 4.13 4.70 (4.40)	24500 (25800) 33000 37050 (39700)
<i>nor</i> Melicopine (IX) .. .. .	(4740) ? 4310 3140 2755 2505	(3.30) ? 3.71 4.23 4.45 4.43	(21100) ? 23250 31900 36300 39950
Melicopidine (VIII) .. .. .	3990 (3170) ?? 2770	3.89 (3.66) ?? 4.66	25100 (31600) ?? 36100
<i>nor</i> Melicopidine (X) .. .. .	4170 (3350) ? (3080) 2830 (2600) ??	3.76 (3.60) ? (4.06) 4.57 (4.38) ??	24000 (29900) ? (32500) 35350 (38500) ??
Melicopicine (V) .. .. .	4030 (3000) 2690	3.88 (3.82) 4.73	24800 (33300) 37250
<i>nor</i> Melicopicine (VI) .. .. .	4210 3120 2750 (2520)	3.76 4.02 4.63 (4.36)	23800 32100 36400 (39700)
Dihydroacronycine (XVII) .. .. .	3845 (3150) (2900) 2710	3.82 (4.09) (4.36) 4.60	26000 (31700) (34500) 36900

TABLE 3 (*Continued*)

Compound	Wavelength	Log $\epsilon$	Frequency
2,3-Dimethoxy-10-methylacridone-1,4-quinone (XIII) .. .. .	(5510)	(2.82)	(18150)
	(4600) ??	(3.27) ??	(21700) ??
	3650	3.92	27400
	3280	3.88	30500
	2740	4.17	36500
	2360	4.46	42350
1,3-Diacetoxy-2-methoxy-4-hydroxy-10-methylacridone (XII) .. .	4120	3.81	24250
	3070	4.03	32550
	2720	4.68	36800
Monobasic acronycinic acid (XVI) ..	3800	3.91	26300
	(3160)	(3.75)	(31650)
	2925	4.16	34200
	2650	4.72	37700
	2270	4.17	44100
Bromacronycine (XX) .. .	3945	3.86	25400
	3025	4.51	33100
	2890	4.52	34600
	2645	4.54	37800
	2290	4.26	43700
Bromnoracronycine (XXI) .. .	4110	3.70	24350
	3010	4.55	33200
	2820	4.54	35450
	(2630)	(4.39)	(38000)
Acetylnormelicopidine (XI) .. .	3995	3.89	25000
	(3250)	(3.61)	(30750)
	2775	4.67	36000
norDihydroacronycine (XIX) .. .	3960	3.76	25300
	3320	4.04	30100
	3000	4.08	33300
	2765	4.65	36200
	2520	4.40	39700
	2290	4.22	43700
2,4-Dihydroxy-10-methylacridone ..	3935	3.81	25400
	3270	3.89	30600
	2940	4.12	34000
	2710	4.68	36900
	2625	4.62	38100
	(2480)	(4.50)	(40300)
2-Methoxy-3-hydroxy-10-methylacridone-1,4-quinone (XIV) .. .	3660	4.01	27300
	3300	3.96	30300
	2770	4.17	36100
	2350	4.44	42550

bands. From comparison with the spectrum of acronycine (the  $V_{2p}$  band of which is very similar in position to that of *nora*acronycine), it is clear that the shift in the  $V_{2p}$  band is greater than that due to the one double bond in the pyran ring so it must be concluded that the side chain in acronycinol is more extensively conjugated. The structure suggested above is the only possibility fulfilling this requirement.

### III. EXPERIMENTAL

The compounds employed for determinations of the spectra were mostly samples purified for elementary analysis. Those which were not analysed samples were carefully purified by standard methods.

The spectra were determined in ethanol using a Spekker ultraviolet spectrophotometer and a Hilger medium quartz spectrograph. The light source was a tungsten-steel spark.

The extinction coefficients apply to cell lengths in cm. and concentrations in mol. litre<sup>-1</sup>.

### IV. ACKNOWLEDGMENTS

The authors would like to express their thanks to Dr. J. R. Price, C.S.I.R.O., for specimens of melicopine, melicopidine, melicopicine, and their derivatives; to Mr. G. K. Hughes, University of Sydney, for samples of evoxanthine, *nor*-evoxanthine, and the monomethoxy- and hydroxy-10-methylacridones; and to Mr. N. L. Lottkowitz for assistance in the preparation of solutions for spectral analysis.

### V. REFERENCES

- (1) BROWN, R. D., DRUMMOND, L. J., LAHEY, F. N., and THOMAS, W. C.—*Aust. J. Sci. Res. A* **2**: 624 (1949).
- (2) CROW, W. D., and PRICE, J. R.—*Aust. J. Sci. Res. A* **2**: 282 (1949).
- (3) HUGHES, G. K., and NEILL, K. G.—*Aust. J. Sci. Res. A* **2**: 429.
- (4) MAYNEORD, W. V., and ROE, E. M. F.—*Proc. Roy. Soc. A* **152**: 311 (1935).
- (5) RADULESCU, D., and OSTROGOVICH, G.—*Ber. dtsh. chem. Ges.* **64**: 2233 (1931).
- (6) COULSON, C. A.—*Proc. Phys. Soc.* **60**: 257 (1948).
- (7) JONES, R. N.—*J. Amer. Chem. Soc.* **67**: 2129 (1945).
- (8) COULSON, C. A., and LONGUET-HIGGINS, H. C.—*Proc. Roy. Soc. A* **192**: 25 (1947).
- (9) LONGUET-HIGGINS, H. C., and COULSON, C. A.—*Trans. Faraday Soc.* **43**: 87 (1947).
- (10) ALBERT, A., and SHORT, L. N.—*J. Chem. Soc.* **1945**: 760 (1945).
- (11) MULLIKEN, R. S., RIEKE, C. A., and BROWN, W. G.—*J. Amer. Chem. Soc.* **63**: 54 (1941).
- (12) HERZBERG, G.—“Infrared and Raman Spectra of Polyatomic Molecules.” (Van Nostrand: New York, 1945.)
- (13) KOHLRAUSCH.—“Ramanspektren.” (1943.)
- (14) THOMPSON, H. W.—*J. Chem. Soc.* **1948**: 328 (1948).
- (15) NITZSCHE, S.—*Angew. Chem.* **52**: 517 (1939).
- (16) PLATTNER, PL. A.—*Helv. Chim. Acta* **24**: 283E (1941).
- (17) PLATTNER, PL. A., and HEILBRONNER, E.—*Helv. Chim. Acta* **31**: 804 (1948).
- (18) BROWN, R. D.—*Trans. Faraday Soc.* **44**: 984 (1948).

# THE ULTRAVIOLET ABSORPTION SPECTRA OF THE ACRIDONE ALKALOIDS

## II. COMPOUNDS RELATED TO 4-QUINOLONE

By R. D. BROWN\* and F. N. LAHEY†

[*Manuscript received June 21, 1950*]

### *Summary*

A qualitative discussion, based mainly upon the MO approximation, is presented for the ultraviolet spectra of naphthalene, quinolone, and some degradation products of the acridone alkaloids related to 4-quinolone.

## I. INTRODUCTION

In Part I(1) the spectra of various degradation products of the alkaloids acronycine, melicopine, evoxanthine, and related alkaloids were discussed in relation to their structures as deduced from purely chemical evidence. The present discussion is concerned with some simpler derivatives of these alkaloids, and the spectra obtained amply confirm the structures attributed to them on chemical grounds. The qualitative interpretations are based almost entirely upon the concepts of molecular orbital theory of quantum mechanics, an approach which seems to have been almost entirely neglected in the past when considering the spectra of large organic molecules, except in the case of some polynuclear aromatic hydrocarbons(2).

## II. GENERAL DISCUSSION OF THE SPECTRA

We may start by comparing the absorption spectra of naphthalene and quinoline(3, p. 363). Except for the longest wavelength region beyond 2900 Å the two curves show bands very similar in position and intensity, the quinoline spectrum being of slightly higher intensity. On theoretical grounds the replacement of a carbon atom in a conjugated system by a nitrogen atom would be expected to disturb the  $\pi$ -electron distribution considerably due to the disparity in the respective Coulomb integrals of the two atoms. Such a perturbation will inevitably alter the energies of the various electronic energy levels of the molecule. In view of the above similarity, the various energy levels participating in the absorption spectrum must be perturbed approximately the same amount in quinoline. A similar phenomenon has already been observed in the case of anthracene, acridine, and phenazine(1), and is borne out by first order perturbation calculations for the energy levels of naphthalene computed by the simplest LCAO molecular orbital approximation (R. D. Brown, unpublished data).

\* Organic Chemistry Laboratory, University of Melbourne.

† Present address: Department of Chemistry, University of Queensland.



The near ultraviolet absorption spectrum of naphthalene has for some time defied adequate interpretation. Uncertainty arose as to the nature of the longest wavelength absorption band when its position, as computed by the simplest HLSP approximation, came to be compared with the experimental position (see, for example, 4). Generally it has been assumed that it consists of two band systems(2), but when it is attempted to relate the spectra of some quinoline derivatives to that of naphthalene it seems necessary to assume the presence of three band systems (*vide infra*). Also the appearance of the longer wavelength band is that of a high-intensity transition superimposed upon one of lower intensity, the latter undergoing a considerable increase in intensity in quinoline, suggesting that it is a symmetry-forbidden transition (e.g.  ${}^1A_{1g} \rightarrow {}^1A_{1g}$ ).

On the theoretical side the simple LCAO treatment(2) indicates that the transitions from the ground state to the two lowest excited singlet levels are symmetry-permitted, the first forbidden transition ( $A_{1g} \rightarrow B_{3g}$ ) being of appreciably greater energy. A more elaborate molecular orbital treatment(5), taking into account electronic repulsion and the very important configurational interaction, inverts the order of the two lowest levels, the first forbidden transition, now  ${}^1A_{1g} \rightarrow {}^1A_{1g}$ , again requiring greater energy. The HLSP treatment including polar structures on the other hand predicts the longest wavelength band to be a forbidden transition and the next two to be  $z$ -polarized\* ( $A_{1g} \rightarrow B_{1u}$ ) and  $y$ -polarized ( $A_{1g} \rightarrow B_{2u}$ ) respectively(5, see Craig's results), and thus seems to be in better agreement with experiment. The disagreement with the more elaborate MO treatment may possibly be due to the difficulty of ensuring that all important configurations have been included when determining the configurational interaction(6) but it seems that the symmetries of the three lowest-lying excited states of naphthalene are still not definitely established theoretically. For the ensuing discussion we shall term the three bands in the naphthalene the  $V_1$ ,  $V_2$ , and  $V_3$  bands respectively in order of increasing frequency.

The ultraviolet absorption spectra of a number of substances related to naphthalene and quinoline, including all of the hydroxyquinolines, have recently been published(7). For all the compounds studied (more than 20) the spectra can be divided into two classes depending upon whether the substituent is in the  $\alpha$ -position or the  $\beta$ -position of naphthalene. In the former class we have compounds such as  $\alpha$ -naphthol, quinoline, 4-hydroxyquinoline, 5-, and 8-hydroxyquinolines, and in the latter  $\beta$ -naphthol, *isoquinoline*, and 7-hydroxy*isoquinoline*. Let us consider these in turn.

It has been mentioned above that the positions of the  $V_1$ ,  $V_2$ , and  $V_3$  bands of quinoline are very similar in position to those of the corresponding bands in the spectrum of naphthalene; if anything they lie at slightly longer wavelengths, and it is thus convenient to compare the spectra of members of this first class with the quinoline spectrum.† The spectrum of  $\alpha$ -naphthol is very

\* The coordinate system employed here is that customarily used(2, 5) and analogous to that used for anthracene(1).

† The fact that the low intensity  $V_1$  band of naphthalene undergoes a great increase in intensity in the spectrum of quinoline is strong evidence that in the latter it is perturbation-permitted(1) and hence in naphthalene it is symmetry-forbidden ( ${}^1A_{1g} \rightarrow {}^1A_{1g}$  or  ${}^1A_{1g} \rightarrow {}^1B_{3g}$ ).

similar to that of quinoline in outline, but all the bands, particularly the  $V_2$  and  $V_3$  bands, are shifted to longer wavelengths. These shifts may be attributed to the delocalization of the oxygen  $2p$  lone-pair electrons over the naphthalene framework, together with another effect due to the very large Coulomb integral of oxygen. The MO calculations of Sandorfy(8) indicate that the combination of these two effects causes a shift of the bands to longer wavelengths. This shift has been observed for many hydroxy compounds (see, for example, 9) and is also found in the present case.

In the compounds 4-, 5-, and 8-hydroxyquinoline again all bands shift to longer wavelengths compared with quinoline, so much so that presumably the small inflections appearing on the shorter wavelength side of the  $V_1$  band (now strongly perturbation-permitted) are due to the presence of the now lower intensity  $V_2$  band. The intensity of the latter band seems to remain about the same as in quinoline for all of these compounds. On the other hand the spectrum of 4-hydroxyquinoline exhibits a very large increase in intensity of the perturbation-permitted  $V_1$  band. A similar increase in intensity, not quite as great, occurs in the spectrum of 2-hydroxyquinoline (*vide infra*) and this may be attributed to the quinolone type of structure causing a greater perturbation of the electron cloud.

Turning now to the second class, the spectrum of *isoquinoline* differs appreciably from that of quinoline. The  $V_1$  band of the former is at slightly longer wavelengths, but the most noticeable feature is the marked shift of the  $V_2$  band to shorter wavelengths in the *isoquinoline* spectrum. These two shifts together produce a resolution of the two bands so that three separate band systems may clearly be seen in the spectrum. This appearance is characteristic of this class of  $\beta$ -substituted naphthalenes.

A comparison of the spectra of 2-, 3-, 6-, and 7-hydroxyquinoline with that of quinoline itself reveals the following common changes. The  $V_1$  band shifts to longer wavelengths while the  $V_2$  band shifts in the opposite direction and suffers a decrease in intensity (except in the case of the 2-isomer which again may be connected with the possibility of attaining a stable quinolone type structure). The  $V_3$  band seems to shift slightly to longer wavelengths but its position in quinoline is uncertain as it lies at the limit of the field of the instruments employed in this work. Again these shifts have the result that the three band systems become well-separated from one another.

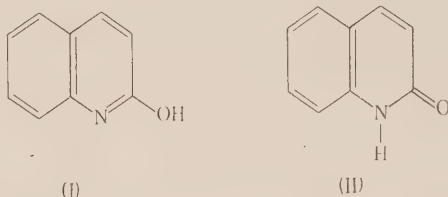
The spectrum of  $\beta$ -naphthol bears the same relation to that of *isoquinoline* as the spectrum of  $\alpha$ -naphthol does to quinoline, but the shift to longer wavelengths is smallest for the  $V_2$  band. The spectrum of 7-hydroxy*isoquinoline* shows a shift of the  $V_1$  band to longer wavelengths and a shift of the  $V_2$  to slightly shorter wavelengths when compared with the spectrum of *isoquinoline*, and similarly the spectra of several other derivatives, without exception, fit into the above classification.

From the above it is apparent that the position of the  $V_2$  band in the spectra of substances related to the naphthalene skeleton is a helpful criterion for the position of substituents.  $\alpha$ -Substitution causes the  $V_2$  band to shift to longer wavelengths, and sometimes to be masked by the  $V_1$  band, while  $\beta$ -substitution

produces a smaller shift in the opposite direction. These shifts would be more easily understandable if the  $V_2$  band consists of a  ${}^1A_{1g}-{}^1B_{2u}$ ; it should also be borne in mind that this has only been investigated for hydroxy-type substituents.

The spectra of 2-hydroxyquinoline, 4-hydroxyquinoline, and their corresponding *N*-methylquinolones are shown in Figure 1. The spectra of 2-hydroxyquinoline (carbostyryl) and its *N*-methyl derivative have been investigated several times before(10-13), mainly with the aim of showing that carbostyryl exists in the quinolone form; the spectrum of 4-hydroxyquinoline has also been recorded(7).

In the case of carbostyryl it was found that its spectrum is very similar to that of the *N*-methyl compound (compare Fig. 1) and differs somewhat from the spectrum of 2-methoxyquinoline. It is worthwhile considering this evidence more fully than any of the previous authors have done. The two structures



under consideration for carbostyryl are the hydroxyquinoline "structure" (I), and the quinolone "structure" (II). In previous discussions it has been assumed that *N*-methylation forces the molecule to assume the structure (II) while *O*-methylation fixes the structure I. However, considering firstly the *N*-methyl compound, there is some evidence(14, 15) that the three N-C bonds in substituted ammonias of the triphenylamine type are coplanar, and this is even more likely when the nitrogen forms part of a ring system. This means that the three  $\sigma$ -bonds are formed by the overlapping of trigonal  $sp^2$  hybridized orbitals, the nitrogen lone pair thus being  $2p$  in nature. The  $2p$  character of these electrons means that they can become delocalized to a much greater extent than  $2s$  lone pairs, and this in turn means that the two C-N bonds in the quinoline ring system have a bond order greater than unity. Also, in view of the fact that the C-O of the quinolone structure is conjugated with the remainder of the ring system, the order of this bond will be less than 2.

Turning now to the *O*-methyl isomer, the oxygen atom here possesses a lone pair occupying  $2p$  atomic orbitals and these become delocalized in the final molecule, in spite of the presence of the methyl group. Thus the order of the C-O bond is greater than unity and the 2-methoxy compound does not possess exactly the structure I. In fact, for the purposes of theoretical calculations of the electronic structures of the *N*- and *O*-methyl derivatives, the difference is caused solely by the very slight difference in Coulomb integrals of O and N due to replacing H by  $\text{CH}_3$  or *vice versa*, and the equally slight effect of hyperconjugation involving the methyl group(16).\* Consequently it is not

\* Other minor effects are due to the lone pair electrons not possessing the pure  $p$  antisymmetry and to the O-Me bond not being held rigidly coplanar with the quinolone ring.

surprising that the ability to discriminate between the spectra of the *O*- and *N*-methyl ethers is the exception rather than the rule(11, 12). In the present case there is a detectable difference in the spectra and this will inevitably mean that the structure of the *N*-methyl ether is nearer to II, of the *O*-methyl ether

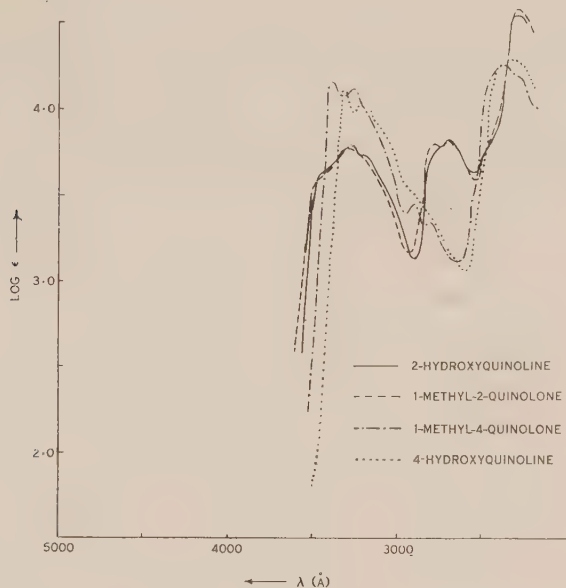
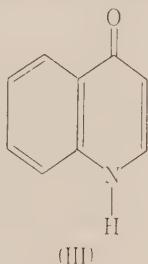


Fig. 1

nearer to I. This difference in structure would be enhanced if the nitrogen and/or oxygen substituent were not coplanar with the quinoline ring system. The fact that carbostyryl has a spectrum closer to its *N*-methyl derivative indicates that its structure also is nearer to II.

4-Hydroxyquinoline provides another case where the spectra of the two methyl ethers, when compared with that of the hydroxy compound, indicate that the electron distribution in the latter tends towards the extreme structure



III. The spectra of 4-hydroxyquinoline and 1-methyl-4-quinolone (Fig. 1) are very similar, the bands of the methyl derivative showing a very slight shift to longer wavelengths, probably to be attributed to hyperconjugation. The spectrum of 4-methoxyquinoline(7) is quite different from that of the parent



hydroxy compound, all bands, and especially the  $V_1$  band, having shifted to much shorter wavelengths in the spectrum of the former.

A comparison of the absorption curves of carbostyryl and 1-methyl-2-quinolone (Fig. 1) with those of the previous workers reveals that the resolution obtained in the present investigation is superior to that of the earlier investigations.

A study of the spectra of some derivatives of 1-methyl-4-quinolone produces some interesting features. Substitution of a bromine atom in position 3 causes all absorption bands to appear at longer wavelengths with a simultaneous increase in all intensities (Fig. 2). These effects would have been anticipated because the bromine atom provides lone pair electrons with the requisite

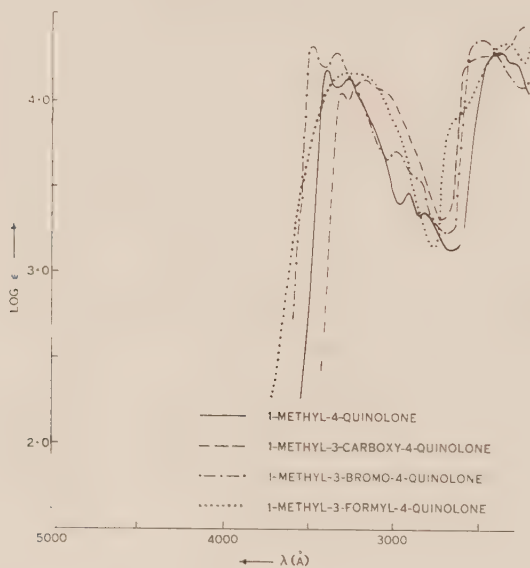


Fig. 2

symmetry properties to become  $\pi$ -electrons; this in general will result in a decreased energy of transition resulting in longer wavelength absorption.\* The increased intensity of absorption arises from a larger transition moment, this latter being a frequent consequence of the introduction of a polar group, such as the bromo group, into a molecule.

On the other hand, when the group substituted in the 3 position is the carboxyl or formyl group the bands shift to shorter wavelengths (Fig. 2). This applies particularly to the  $V_1$  band system, which occurs in the region of 3000  $\text{\AA}$ . The shift in the longest wavelength band is at first sight unexpected. A comparison of the spectra of benzene and benzoic acid, benzene and benzaldehyde(3, pp. 365, 372) or even ethylene and acrolein, shows that both carboxyl and formyl groups produce a shift to longer wavelength, this being

\* Note that this position has a high local  $\pi$ -electron density similar to positions 1 and 3 in acridone (cf. 1).

attributed theoretically to the increased conjugated system. It therefore seems that in the cases under consideration the quinolone  $\pi$ -orbitals do not extend appreciably over the substituents. Further evidence for this is provided by consideration of the intensities of the band systems. It was mentioned in connection with the spectrum of 1-methyl-3-bromo-4-quinolone that the introduction of a polar group *as part of the conjugated system* causes an increase in intensity of the bands but in the spectra of the formyl and carboxy derivatives the intensities are little altered.

The absence of conjugation is readily attributable to steric hindrance. Fisher-Hirschfelder models of the two molecules indicate the close proximity of the substituents to the carbonyl oxygen. Jones(9) came to the conclusion that

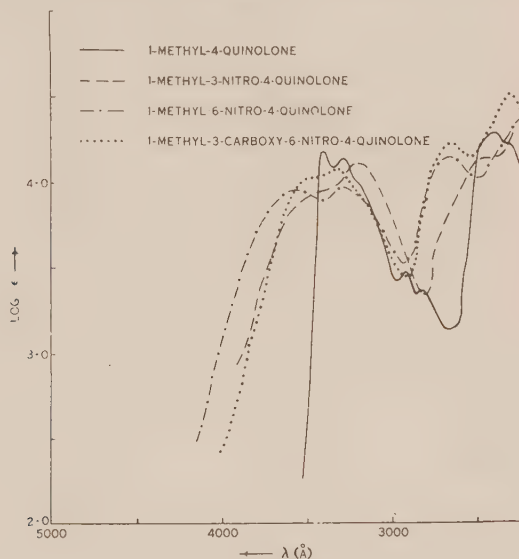


Fig. 3

such molecular models are not sufficiently accurate to predict the onset of steric effects preventing conjugation: in a study of the spectra of derivatives of polynuclear hydrocarbons he noticed that steric interactions were very prevalent in formyl and carboxyl derivatives, and it seems that the van der Waals radius of the formyl group in particular is much greater than indicated by the models.

In the case of the corresponding 3-nitro derivative (Fig. 3) the  $V_1$  band system is very extended, a smaller value being found for the maximum absorption intensity. As the nitro group must suffer considerable steric interaction with the quinolone oxygen, conjugation will be prevented and, in analogy with the carboxy and formyl derivatives, a shift to shorter wavelengths would be anticipated.\* However, the nitro group is known to have unusual effects upon

\* In the absence of conjugation, which will tend to shift the band system to longer wavelengths, the effect upon the electronic structure will be due mainly to a change in the Coulomb integral of the carbon atom at the 3 position. This could conceivably shift the band system in either direction, depending upon the particular case involved.

TABLE 1  
 COORDINATES OF ABSORPTION MAXIMA

The wavelengths,  $\lambda$ , are in Angstrom units, the frequencies,  $\nu$ , in  $\text{cm}^{-1}$ . Data applying to points of inflection are given in parentheses, and indefinite inflections are denoted by ?.

Compound	$\lambda$	$\text{Log } \epsilon$	$\nu$
2-Hydroxyquinoline .. ..	(3400)	(3.66)	(29400)
	3290	3.80	30400
	(3220)	(3.74)	(31100)
	(2760)	(3.78)	(36300)
	2695	3.85	37200
	2300	4.57	43500
1-Methyl-2-quinolone .. ..	(3440)	(3.61)	(29100)
	3300	3.77	30400
	2790	3.81	35900
	2710	3.82	36900
	2300	4.61	43500
4-Hydroxyquinoline .. ..	3325	4.16	30100
	3195	4.10	31300
	2330	4.35	42900
1-Methyl-4-quinolone .. ..	3400	4.17	29400
	3260	4.14	30700
	2910	3.47	34400
	2810	3.36	35600
	2390	4.30	41800
	(2330)	(4.24)	(42900)
3-Carboxy-4-hydroxyquinoline ..	(3220)	(3.98)	(31100)
	3100	4.11	32300
	(3030)	(4.04)	(33000)
	2270	4.43	44100
1-Methyl-3-carboxy-4-quinolone ..	3290	4.05	30400
	3160	4.13	31700
	(3080)	(4.09)	(32500)
	(2480)	(4.27)	(40300)
1-Methyl-3-formyl-4-quinolone ..	3250	4.17	30800
	(2640)	(3.90)	(37900)
	2330	4.35	42900
1-Methyl-3-bromo-4-quinolone ..	3470	4.32	28800
	3225	4.28	30100
	2985	3.71	33600
	(2860)	(3.55)	(35000)
	2500	4.36	40000

TABLE I (Continued)

Compound	$\lambda$	Log $\epsilon$	$\bar{\nu}$
1-Methyl-3-nitro-4-quinolone ..	(3440) ?	(3.94) ?	(29100) ?
	3190	4.11	31400
	2440	4.15	41000
1-Methyl-6-nitro-4-quinolone ...	3560	3.95	28100
	3280	3.98	30500
	2670	4.16	37500
	2245	4.38	44600
1-Methyl-3-carboxy-6-nitro-4-quinolone	3450	4.05	29000
	3310	4.08	30200
	2645	4.23	37900
	2290	4.53	43700
1-Methyl-8-methoxy-4-quinolone ..	3400	4.05	29400
	3310	4.07	30200
	(3080) ?	(3.88) ?	(32500) ?
	2600	3.76	38500
	2275	4.52	44000
Acid I (N/100 sodium hydroxide)	3370	4.07	29700
	3270	4.08	30600
	2465	4.52	40600
Acid II (N/100 sodium hydroxide)	4510	3.92	22200
	(2900) ?	(3.78) ?	(34500) ?
1-Methyl-3-carboxy-5, 6, 7, 8-tetrahydro-4-quinolone	(2880)	(3.53)	(34700)
	2625	3.96	38100
	2200 ?	4.46	45500 ?

the absorption spectrum(17). In particular it is found that although the nitro group alone has an absorption very similar to the carbonyl group(17, 18), in conjugated systems such as 1-nitrobutadiene or  $\beta$ -nitrostyrene a new high intensity band (log  $\epsilon$  more than 4) appears around 3000 Å, which has no parallel in the analogous carbonyl compounds. The absorption of an isolated nitro group is presumably due to a transition involving the excitation of a  $2p_y$  lone-pair electron of one of the oxygen atoms to an excited  $\pi$ -orbital of the N-O bond, this transition being forbidden by the symmetry selection rules just as in the corresponding carbonyl transition. The divergence from the theoretically well-understood carbonyl behaviour in conjugated molecules indicates that the "new" high intensity "nitro-band" is not directly connected with delocalization of the  $\pi$ -electrons of the nitro group because it is this analogous delocalization which produces the (quite different) changes in the spectra of carbonyl compounds. Thus it does not necessarily follow that because conjugation of the nitro group with the quinolone system is inhibited by steric interactions the



possibility of the appearance of the nitro-band is ruled out; indeed the spectra of the three nitro derivatives of 4-quinolone cannot easily be explained unless the appearance of this nitro-band is assumed.

Consequently it is probable that the longest wavelength band in 3-nitro-1-methyl-4-quinolone consists of two bands, one being the quinolone  $V_1$  band displaced towards shorter wavelengths (corresponding to the maximum at 3190 Å) and the other (corresponding to the inflection around 3440 Å) being the new band attributable to the nitro group itself, the "nitro-band". That the longest wavelength band system is in reality two electronic bands seems likely from the great width of the absorption band together with its relatively high intensity: if this were due to just one electronic transition the corresponding oscillator strength would be unusually high.

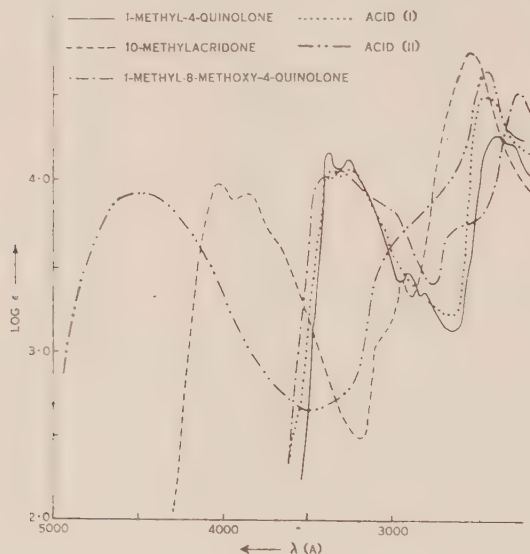


Fig. 4

Another minor point to notice is that the presence of the nitro group in the 3-position has caused the  $V_2$  band to shift to the ultraviolet. This together with the evidence of prevention of conjugation of the nitro group with the quinolone system (shift of the  $V_1$  band system in the same direction) is evidence of a purely spectrographical nature that the nitro group is not in the benz-ring; the latter has been established on chemical grounds(19, 20).

The longest wavelength band of the 6-nitro derivative also is clearly a combination of two band systems. Again the  $V_2$  band is shifted to the ultraviolet (Fig. 3). The longest wavelength band at 3560 Å corresponds to the  $V_1$  band of the quinolone spectrum, displaced in this direction by conjugation of the 6-nitro group with the 4-quinolone ring system; in this case there is no steric interaction to prevent this conjugation. The second maximum at 3280 Å is in the position expected for the nitro-band, and the maximum at 2245 Å is presumably the quinolone  $V_3$  band.

The absorption spectrum of 1-methyl-3-carboxy-6-nitro-4-quinolone is shown in Figure 3. The band at 3450 Å is the  $V_1$  band, shifted to longer wavelengths due to the nitro group participating in the conjugation. The second maximum at 3310 Å seems to be the nitro-band and the other bands at 2645 and 2290 Å correspond to the  $V_2$  and  $V_3$  bands. The position of the  $V_2$  band is very close to that of 6-nitro-1-methyl-4-quinolone, in each case the shift to shorter wavelengths being predominantly due to the  $\beta$ -nitro group.

The spectrum of 8-methoxy-1-methyl-4-quinolone (Fig. 4) shows the quinolone  $V_1$  band practically unaltered in position. This, as discussed above for 3-substituents, is to be attributed to steric inhibition of conjugation. In this case the conjugation is due to delocalization of the  $2p$  lone pair electrons on the methoxyl oxygen, which will only be possible if the O-Me bond lies in or near the plane of the quinolone ring.

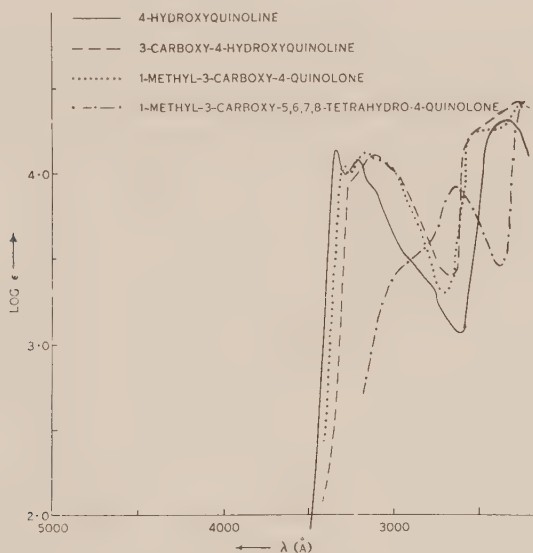
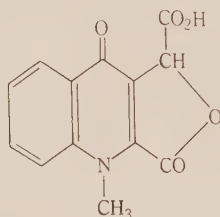


Fig. 5

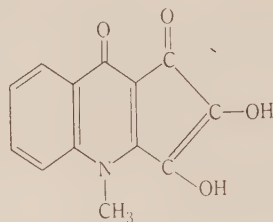
The spectrum of 3-carboxy-4-hydroxyquinoline (Fig. 5) again exhibits this phenomenon of steric inhibition of conjugation due to interaction of the carboxyl and hydroxyl groups. A comparison of this spectrum with the spectrum of the corresponding *N*-methyl acid reveals the same close similarity as do the spectra of 4-hydroxyquinoline and 1-methyl-4-quinolone, again bands of the *N*-methyl derivative appearing at slightly longer wavelengths due probably to hyperconjugation.

Reduction of 1-methyl-3-carboxy-4-quinolone to the 5,6,7,8-tetrahydro acid produces, as expected, a considerable shift of the bands to shorter wavelengths, and the spectrum, as anticipated, is similar to that of 1-methyl-4-pyridone(19, 20). The position of the maximum (2625 Å) is the same as that recorded by Specker and Gawrosch for the latter compound in methanol, but the intensity is less ( $\log \epsilon$  3.96 as opposed to 4.26).

Finally, in the quinolone series we come to two compounds whose structures are not quite so well established as the other compounds so far studied(21). These will be termed for convenience acid I and acid II and their probable structures are shown in the accompanying diagrams. The ultraviolet absorption



acid (I)



acid (II)

spectrum of acid I in N/100 sodium hydroxide is shown in Figure 4. It is seen to resemble the spectrum of 1-methyl-4-quinolone quite closely. The  $V_1$  band especially is very similar except that its intensity is a little less than the  $V_1$  band of 1-methyl-4-quinolone. The spectrum thus suggests that acid I is a derivative of 1-methyl-4-quinolone and that the substituents are such that any conjugation with the quinolone system is precluded by steric or other effects. Clearly the proposed formula fulfils these requirements.

The absorption spectrum of acid II, also in N/100 sodium hydroxide (Fig. 4), is completely different from that of 1-methyl-4-quinolone and in general shape is more reminiscent of the acridone spectrum(1). However, its long wavelength band is at considerably longer wavelengths than any of the acridone derivatives so far examined apart from the quinones, so the spectrum suggests a conjugated structure of larger size than quinolone, or alternatively with an unusual arrangement of bonds.\* The structure proposed is quite likely to give rise to such a spectrum. The structure is that of a quinolone derivative of cyclopentadienone, the known aromatic derivatives of this system (see, for example, 22) being coloured, and the hydroxy groups introduce the possibility of keto-enol tautomerism (the alkaline solution has the same colour to the eye as the solid acid). The keto form is a diketone and these are frequently characterized by visible colour. These considerations indicate that at least there is no spectrographic evidence at conflict with the proposed structure.

### III. CONCLUSION

In the preceding discussion an attempt has been made to give a plausible explanation of the spectra of the various molecules, based as far as possible on the qualitative concepts of the quantum-mechanical molecular orbital approximation. The discussion is ultimately based upon the spectrum of naphthalene. Unfortunately it is somewhat speculative because of the uncertainty in the nature of the lowest excited valence states of naphthalene. The spectrum of the latter cannot be clarified by comparison with other spectra

\* An unusual arrangement of bonds, such as in the molecules fulvene and azulene, is frequently characterized by light absorption in the visible region.

because no such suitable compounds exist, benzene being too symmetrical, and a comparison with anthracene apparently not leading to the correct conclusions. On the theoretical side, as mentioned in the discussion, no final decision has been reached although the VB treatment is the most attractive.

The present experimental study suggests the following interpretations: (a) the  $V_1$  band of naphthalene is due to a forbidden transition (upper state  $A_{1g}$  or  $B_{3g}$ ); (b) the  $V_2$  band is due to a (permitted)  $y$ -polarized transition (upper state  $B_{2u}$ ), the evidence being the effect of  $\alpha$ - and  $\beta$ -substitution upon the band position; (c) the  $V_3$  band, being usually at the short wavelength limit of the instrument employed in the present study, has not been sufficiently investigated to suggest its polarization; it is certainly a permitted transition so the possible symmetries for the upper level are  ${}^1B_{1u}$ ,  ${}^1B_{2u}$ . Further work may serve to illuminate the latter designation and to provide further evidence for interpretation of the  $V_2$  band.

#### IV. EXPERIMENTAL

The instruments etc. were the same as those used for the compounds described in Part I. The solvent for the spectra determined in N/100 sodium hydroxide was prepared from conductance water and analar sodium hydroxide. Except where otherwise indicated all spectra were determined in purified ethanol.

#### V. ACKNOWLEDGMENTS

The authors are grateful to Dr. J. R. Price, C.S.I.R.O., for providing samples of a number of the quinolone derivatives studied in the present investigation, and to Mr. N. L. Lottkowitz for assistance in the preparation of solutions for spectral analysis. They are also grateful to Mr. A. N. Hambly for his helpful suggestions during preparation of the manuscript.

#### VI. REFERENCES

- (1) BROWN, R. D., and LAHEY, F. N.—*Aust. J. Sci. Res. A* **3**: 593 (1950).
- (2) COULSON, C. A.—*Proc. Phys. Soc.* **60**: 257 (1948).
- (3) WASHBURN, E. D. (Editor).—*International Critical Tables*, Vol. 5 (1929).
- (4) MACCOLL, A.—*Quart. Rev.* **1**: 54 (1947).
- (5) JACOBS, J.—*Proc. Phys. Soc.* **62**: 710 (1949).
- (6) CRAIG, D. P.—*Proc. Roy. Soc. A* **200**: 485 (1950).
- (7) EWING, G. W., and STECK, E. A.—*J. Amer. Chem. Soc.* **68**: 2181 (1946).
- (8) SANDORFY, C.—*C.R. Akad. Sci. Paris* **227**: 1034 (1948).
- (9) JONES, R. N.—*J. Amer. Chem. Soc.* **67**: 2136 (1945).
- (10) HARTLEY, W. N., and DOBBIE, J. J.—*Trans. Chem. Soc.* **75**: 640 (1899).
- (11) MORTON, R. A., and ROGERS, E.—*Trans. Chem. Soc.* **127**: 2698 (1925).
- (12) LEY, H., and SPECKER, H.—*Ber. dtsh. chem. Ges.* **72**: 192 (1939).
- (13) MORTON, R. A., and STUBBS, A. L.—*J. Chem. Soc.* **1939**: 1321 (1939).
- (14) WEIZMANN, A.—*Trans. Faraday Soc.* **36**: 978 (1940).
- (15) BERGMANN, E., and SCHÜTZ, W.—*Z. phys. Chem. B* **19**: 401 (1932).
- (16) MULLIKEN, R. S., RIEKE, C. A., and BROWN, W. G.—*J. Amer. Chem. Soc.* **63**: 54 (1941).
- (17) BRAUDE, E. A., JONES, E. R. H., and ROSE, G. G.—*J. Chem. Soc.*, **1947**: 1104 (1947).
- (18) BRAUDE, E. A.—*Ann. Rep.* **42**: 105 (1945).
- (19) SPECKER, H., and GAWROSCH, H.—*Ber. dtsh. chem. Ges.* **75**: 1344 (1942).
- (20) RIEGEL, R., and REINHARD, M. C.—*J. Amer. Chem. Soc.* **48**: 1334 (1926).
- (21) CROW, W. D., and PRICE, J. R.—*Aust. J. Sci. Res. A* **2**: 282 (1949).
- (22) DILTHEY, W., HENKELS, S., and SCHAEFER, A.—*Ber. dtsh. chem. Ges.* **71**: 974 (1938).



# COLOURING MATTERS OF AUSTRALIAN PLANTS

## I. THE STRUCTURE OF DROSERONE

By R. G. COOKE\* and W. SEGAL\*

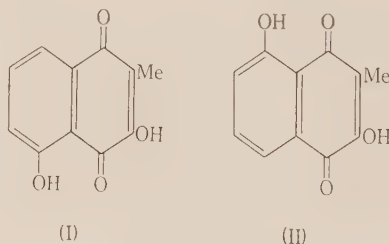
[Manuscript received June 20, 1950]

### Summary

Droserone, one of the pigments occurring in *Drosera whittakeri*, is proved to be 3,5-dihydroxy-2-methyl-1,4-naphthoquinone by a five step synthesis from the known 5-methoxy-1-tetralone. From an intermediate product, 2-hydroxy-5-methoxy-1,4-naphthoquinone, the droserone isomer 2,5-dihydroxy-3-methyl-1,4-naphthoquinone and 2-hydroxyjuglone are also obtained. The latter compound is significant in the chemistry of the natural pigment juglone. An interpretation is given for some orientation effects observed in the nucleophilic substitution of certain naphthoquinones.

### I. INTRODUCTION

Investigation of the pigments of *Drosera whittakeri* by Rennie(1, 2) and by Macbeth and co-workers(3-5) showed that hydroxydroserone has the structure 3,5,8-trihydroxy-2-methyl-1,4-naphthoquinone, and that droserone must be either 3,5- or 3,8-dihydroxy-2-methyl-1,4-naphthoquinone (I or II).



Winzor(6) confirmed the structure of hydroxydroserone by synthesis, and Thomson(7) has shown recently that droserone is 3,5-dihydroxy-2-methyl-1,4-naphthoquinone (I) by synthesis of both the compounds I and II.

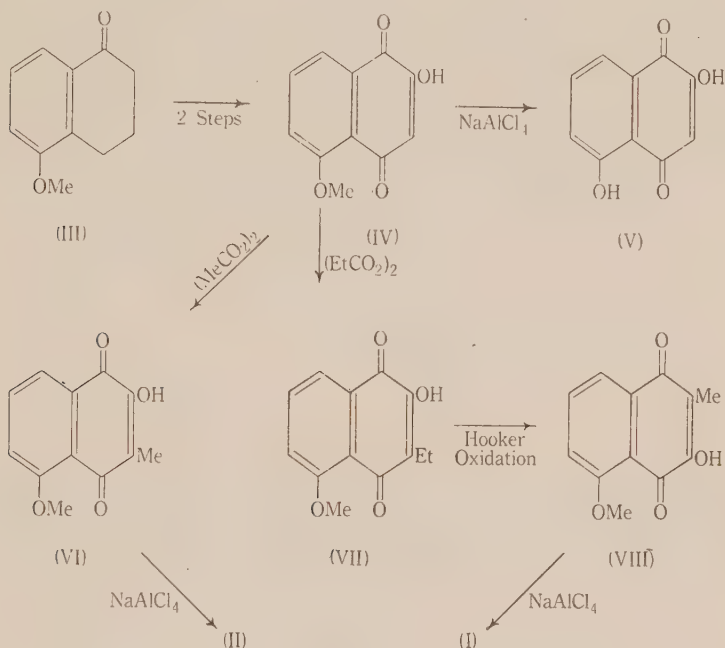
A more direct proof of these structures had been accomplished independently in this Laboratory, and the synthetic work was almost complete when Thomson's paper was published. Details of this alternative synthesis are now presented.

### (a) Synthesis of Droserone

By means of the general method of Buu-Hoi and Cagniant(8), 5-methoxy-1-tetralone (III) is converted into 2-hydroxy-5-methoxy-1,4-naphthoquinone (IV). Alkylation of IV by the method of Fieser and Oxford(9) gives 2-hydroxy-5-methoxy-3-methyl-1,4-naphthoquinone (VI) and 2-hydroxy-5-methoxy-3-ethyl-1,4-

\* Department of Chemistry, University of Melbourne.

*naphthoquinone* (VII). Hooker oxidation(10, 11) of VII gives *3-hydroxy-5-methoxy-2-methyl-1,4-naphthoquinone* (VIII) isomeric with VI. On demethylation IV, VI, and VIII give respectively 2,5-dihydroxy-1,4-naphthoquinone (2-hydroxyjuglone, V), 2,5-dihydroxy-3-methyl-1,4-naphthoquinone (II), and 3,5-dihydroxy-2-methyl-1,4-naphthoquinone (droserone, I).



As shown by Fieser and co-workers(12), the Hooker oxidation not only shortens the alkyl side-chain by one carbon atom, but also reverses the positions of the alkyl and hydroxyl groups on the quinone ring.

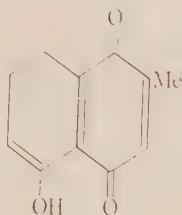
#### (b) Orientation of Substituents in Juglone Derivatives

From a study of absorption spectra and reduction potentials Macbeth and co-workers(5) concluded that droserone is closely related in structure to the hydroxyjuglone obtained by Mylius(13) through the oxidation of juglone (5-hydroxy-1,4-naphthoquinone). The orientation of this hydroxyjuglone was not known, so our synthesis of droserone and its isomer was originally designed to establish the relationship of all these compounds. As shown above 2-hydroxyjuglone can be obtained from the same intermediate, and as it is different from the compound obtained by oxidation of juglone the latter must be 3-hydroxyjuglone. Subsequently Thomson(14) determined the orientation of the hydroxyjuglones by a different method. His 2-hydroxyjuglone is identical with the 2,5-dihydroxy-1,4-naphthoquinone (V) prepared as shown above.

The absorption spectra of synthetic droserone and its isomer (II) are shown in Figure 1. It is evident that in the region examined by Lugg, Macbeth, and

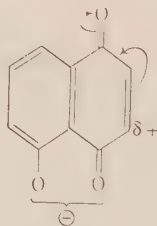
Winzor(5) the light absorption is not sufficiently reliable evidence of the orientation of the hydroxyl groups. However, the visible colour of the solids does give a reliable indication of orientation. In all the isomeric pairs we have examined, the compound having oxygen substituents in the 2,5 positions is invariably deeper in colour than the corresponding 3,5 derivative.

Establishment of the 3,5 orientation of the hydroxyl groups in both droserone and the hydroxyjuglone obtained by oxidation of juglone was not unexpected. It is evident that droserone is a hydroxy derivative of plumbagin (IX) which has been obtained from another sundew (*Drosera rotundifolia* L.) by Witanowski(15) and by Dieterle(16, 17).



(IX)

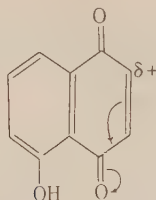
Furthermore, the formation of 3,5-dihydroxy-1,4-naphthoquinone by oxidation of juglone or hydrojuglone in alkaline media(13) is evidently due to nucleophilic hydroxylation of the juglone anion, which may be represented as X.



(X)

Preferential nucleophilic substitution at C<sub>3</sub> could be interpreted as due to activation as shown, since the anionic charge should be distributed over the oxygen atoms at C<sub>4</sub> and C<sub>5</sub>, thus diminishing the electrophilic character of the carbonyl group at C<sub>4</sub>.

Thomson(14) has demonstrated that nucleophilic substitution of juglone by dimethylamine results in preferential attack at C<sub>2</sub>. If it be assumed that in this case the reaction is between the weak base and the juglone molecule, the orientation may be interpreted as due to increased electronegativity of the oxygen atom at C<sub>4</sub> through the influence of the hydrogen bond. This is illustrated in XI.



(XI)

Thomson's observation(14) that similar substitutions occur at C<sub>3</sub> in the acetates of juglone and 2,3-dihalogeno-juglones is consistent with this interpretation. The orientation effect in the acetates would be expected to resemble that of the anion rather than that of the hydroxy compound itself. Further investigation of these orientation effects is in progress.

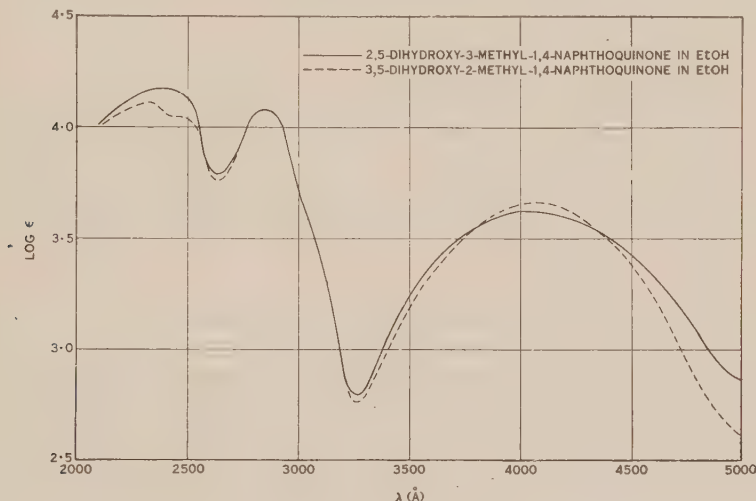


Fig. 1.—Absorption spectra.

## II. EXPERIMENTAL

All melting points are corrected. Methoxyl determinations by N. L. Lottkowitz and assistants; other microanalyses and absorption spectra by Weiler and Strauss, Oxford.

(i) *5-Methoxy-1-Tetralone (III)*.—The method of Mitter and De(18) was used for the preparation of  $\gamma$ -(*o*-hydroxyphenyl)butyric acid which was then methylated with dimethyl sulphate in hot aqueous alkali. The resulting  $\gamma$ -(*o*-anisyl)butyric acid was cyclized as described by Lockett and Short(19) by refluxing for 1 hour with phosphorus oxychloride in tetrachlorethane. After removing the solvent in steam, the tetralone was isolated by prolonged steam distillation. The 5-methoxy-1-tetralone separated from light petroleum in colourless plates, m.p. 89.5–90.5 °C. (lit.(19), 89–89.5 °C.). Semicarbazone, needles from ethanol, m.p. 248–249 °C. (lit.(19), 249–250 °C.).

(ii) *2-Hydroxy-5-Methoxy-1,4-Naphthoquinone (IV)*.—The tetralone (III) was converted to the quinone by the general method of Buu-Hoï and Cagniant(8; see also Brunner and Singule 20). The tetralone (III, 1 g.) and *p*-nitrosodimethylaniline (2 g.) were dissolved in absolute ethanol (60 ml.), and 10% sodium hydroxide (2 ml.) was added. The mixture was allowed to stand at room temperature for 24 hours and then refluxed for 1 hour. Finally



15–20 ml. of solvent was evaporated and the solution was allowed to cool. The *dianil* separated as a dark solid, m.p.  $>165^{\circ}\text{C}$ . Yield 40%. The product crystallized from ethanol in small, violet-black prisms, m.p.  $167\text{--}168^{\circ}\text{C}$ . (decomp.).

Found: C, 73.1; H, 6.4; N, 12.6%.

Calculated for  $\text{C}_{22}\text{H}_{28}\text{O}_2\text{N}_4$ : C, 73.6; H, 6.4; N, 12.7%.

The crystalline dianil was hydrolysed by refluxing for 1 hour with 5% sulphuric acid (50 ml.). The mixture was then extracted with ether in a continuous extractor, and the ethereal solution was shaken repeatedly with 5% aqueous sodium carbonate until the aqueous layer was no longer coloured. After heating the latter to expel ether, the solution of quinone salt was boiled with charcoal, filtered, and acidified while still warm. When quite cold, the granular precipitate was collected and dried. Yield 70%. The product (IV) was sublimed in a high vacuum and crystallized from aqueous ethanol, giving long yellow needles of a *monohydrate* which lost its crystalline form on heating and had m.p.  $176\text{--}177^{\circ}\text{C}$ .

Found: C, 59.7; H, 4.8% (air dried material).

Calculated for  $\text{C}_{11}\text{H}_8\text{O}_4\cdot\text{H}_2\text{O}$ : C, 59.4; H, 4.5%.

After prolonged drying at  $105^{\circ}\text{C}$ .

Found: C, 64.1; H, 4.0; OMe, 15.2%.

Calculated for  $\text{C}_{11}\text{H}_8\text{O}_4$ : C, 64.7; H, 4.0; OMe, 15.2%.

Further quantities of the hydroxyquinone (IV) were usually isolated by hydrolysis of the solution from which the dianil originally separated. The best overall yield of IV from the tetralone was 36%.

The *acetate*, prepared by warming with acetic anhydride and a trace of sulphuric acid, formed yellow prisms from ethanol, m.p.  $140\text{--}141^{\circ}\text{C}$ .

Found: C, 63.4; H, 4.2%.

Calculated for  $\text{C}_{13}\text{H}_{10}\text{O}_5$ : C, 63.4; H, 4.1%.

(iii) *2,5-Dihydroxy-1,4-Naphthoquinone* (V).—A molten mixture of anhydrous aluminium chloride (20 g.) and sodium chloride (4 g.) was cooled to about  $100^{\circ}\text{C}$ . and then added to the methoxyquinone (IV, 0.1 g.). The mixture was heated rapidly to  $180^{\circ}\text{C}$ ., maintained at this temperature for 4–5 minutes and then allowed to cool. After addition of ice and hydrochloric acid, the mixture was allowed to stand overnight and the product was then filtered. After sublimation in a high vacuum, and crystallization from benzene, it formed small orange plates, m.p.  $218^{\circ}\text{C}$ . (decomp., after gradual darkening).

Found: C, 62.7; H, 3.4%.

Calculated for  $\text{C}_{10}\text{H}_6\text{O}_4$ : C, 63.1; H, 3.2%.

The diacetate was prepared by warming with acetic anhydride and a small fragment of anhydrous zinc chloride. It crystallized from ethanol in flat yellow needles, m.p.  $153^{\circ}\text{C}$ ., not depressed by a sample supplied by Dr. R. H. Thomson, and prepared as described previously(14).

Found: C, 60.8; H, 3.7%.

Calculated for  $\text{C}_{14}\text{H}_{10}\text{O}_6$ : C, 61.3; H, 3.7%.

(Thomson(14) did not report any analytical figures for the dihydroxyquinone or the diacetate.)

(iv) *2-Hydroxy-5-Methoxy-3-Methyl-1,4-Naphthoquinone* (VI).—A solution of 2-hydroxy-5-methoxy-1,4-naphthoquinone (IV, 0.2 g.) in glacial acetic acid (3 ml.) was alkylated with diacetyl peroxide (1.1 equivalents) according to the general directions of Fieser and Oxford(9). Yield of product, m.p.  $194\text{--}195^{\circ}\text{C}$ ., was 36%. After crystallization from ethanol it formed orange-yellow needles, m.p.  $196\text{--}197^{\circ}\text{C}$ .

Found: C, 65.9; H, 4.5; OMe, 14.0%.

Calculated for  $\text{C}_{12}\text{H}_{10}\text{O}_4$ : C, 66.0; H, 4.6; OMe, 14.2%.

(v) *2-Hydroxy-5-Methoxy-3-Ethyl-1,4-Naphthoquinone* (VII).—Alkylation of IV with dipropionyl peroxide was effected in the same way. Yield 36%, m.p. 159–162 °C. After sublimation ( $10^{-5}$  mm., 100 °C.) and crystallization from aqueous ethanol, the product formed golden needles, m.p. 162.5–163.5 °C.

Found: C, 67.1; H, 5.4%.

Calculated for  $C_{13}H_{12}O_4$ : C, 67.2; H, 5.2%.

(vi) *3-Hydroxy-5-Methoxy-2-Methyl-1,4-Naphthoquinone* (VIII).—The Hooker oxidation of VII was best effected under the general conditions of the two stage process described by Fieser and Fieser(11). The intermediate ketol was not isolated. The product was sublimed in a high vacuum and crystallized from ethanol, forming fine, matted, pale yellow needles, m.p. 173–174 °C. Yield 70%.

Found: C, 66.2; H, 4.8; OMe, 14.4%.

Calculated for  $C_{12}H_{10}O_4$ : C, 66.0; H, 4.6; OMe, 14.2%.

(vii) *2,5-Dihydroxy-3-Methyl-1,4-Naphthoquinone* (II).—The 5-methyl ether (VI) was demethylated as described above, the mixture being heated to 160 °C. for 6 minutes. After isolation in the usual way, the product was sublimed ( $10^{-6}$  mm., 100–110 °C.) and crystallized from ethanol, forming orange-red prisms, m.p. 192–193 °C. (Thomson(7) reported light orange plates, m.p. 193 °C.) Yield 60%.

Light absorption in ethanol:  $\lambda_{max}$ , m $\mu$ , 237,285,402; log  $\epsilon$ , 4.16, 4.08, 3.62.

Found: C, 64.3; H, 4.1%.

Calculated for  $C_{11}H_8O_4$ : C, 64.7; H, 4.0%.

The diacetate was prepared by warming with acetic anhydride and a trace of sulphuric acid. After crystallization from methanol it formed pale yellow needles, m.p. 163–164 °C. (Thomson(7) reported m.p. 164 °C.)

(viii) *3,5-Dihydroxy-2-Methyl-1,4-Naphthoquinone* (Droserone, I).—The 5-methyl ether was demethylated by the same general method using aluminium chloride and sodium chloride, the melt being heated at 170 °C. for 2–3 minutes. These conditions were evidently too severe and most of the quinone appeared to be resinified after isolation in the usual way. The precipitated solid was very dark in colour, but by adding the material obtained by ether extraction of the filtrate, and subliming the combined solids in vacuum, a small quantity of yellow quinone was obtained. After crystallization from aqueous ethanol containing a little acetic acid, the product formed golden-yellow needles, m.p. 179.5–180.5 °C. (The recorded melting point of natural droserone is 181 °C., but none was available for direct comparison.)

Light absorption in ethanol:  $\lambda_{max}$ , m $\mu$ , 232,285,409; log  $\epsilon$ , 4.12, 4.08, 3.66. Lugg, Macbeth, and Winzor(5) reported  $\lambda_{max}$ , m $\mu$ , 238,410; log  $\epsilon$ , 4.1, 3.7 for the natural pigment.

### III. ACKNOWLEDGMENTS

The authors are indebted to Dr. R. H. Thomson, University of Aberdeen, for samples of 2-hydroxyjuglone and its diacetate, and for personal communications.

### IV. REFERENCES

- (1) RENNIE, E. H.—*J. Chem. Soc.* **51**: 371 (1887).
- (2) RENNIE, E. H.—*J. Chem. Soc.* **63**: 1083 (1893).
- (3) MACBETH, A. K., PRICE, J. R., and WINZOR, F. L.—*J. Chem. Soc.* **1935**: 325 (1935).
- (4) LUGG, J. W. H., MACBETH, A. K., and WINZOR, F. L.—*J. Chem. Soc.* **1936**: 1457 (1936).
- (5) LUGG, J. W. H., MACBETH, A. K., and WINZOR, F. L.—*J. Chem. Soc.* **1937**: 1597 (1937).
- (6) WINZOR, F. L.—*J. Chem. Soc.* **1935**: 336 (1935).
- (7) THOMSON, R. H.—*J. Chem. Soc.* **1949**: 1277 (1949).
- (8) BUU-HOI, and CAGNIANT, P.—*C.R. Akad. Sci. Paris* **214**: 87 (1942).

- (9) FIESER, L. F., and OXFORD, A. E.—*J. Amer. Chem. Soc.* **64** : 2060 (1942).
- (10) HOOKER, S. C.—*J. Amer. Chem. Soc.* **58** : 1168 (1936).
- (11) FIESER, L. F., and FIESER, M.—*J. Amer. Chem. Soc.* **70** : 3215 (1948).
- (12) FIESER, L. F., HARTWELL, J. L., and SELIGMAN, A. M.—*J. Amer. Chem. Soc.* **58** : 1223 (1936).
- (13) MYLIUS, F.—*Ber. deutsch. chem. Ges.* **18** : 463 (1885).
- (14) THOMSON, R. H.—*J. Org. Chem.* **13** : 870 (1948).
- (15) WITANOWSKI, W. R.—*Wiad. farm.* **16** : 420, 432 (1934); *Chem. Abstr.* **28** : 7257 (1934).
- (16) DIETERLE, H.—*Pharm. Arch.* **273** : 235 (1935).
- (17) DIETERLE, H., and KRUTA, E.—*Pharm. Arch.* **274** : 457 (1936).
- (18) MITTER, P. C., and DE, S.—*J. Indian Chem. Soc.* **16** : 35 (1939).
- (19) LOCKETT, J., and SHORT, W. F.—*J. Chem. Soc.* **1939** : 787 (1939).
- (20) BRUNNER, O., and SINGULE, E.—*Monatsh.* **77** : 251 (1947).

# A SYNTHESIS OF SPARTEINE AND SOME RELATED SUBSTANCES\*

By E. F. L. J. ANET,<sup>†</sup> G. K. HUGHES,<sup>†</sup> and E. RITCHIE<sup>†</sup>

[Manuscript received August 2, 1950]

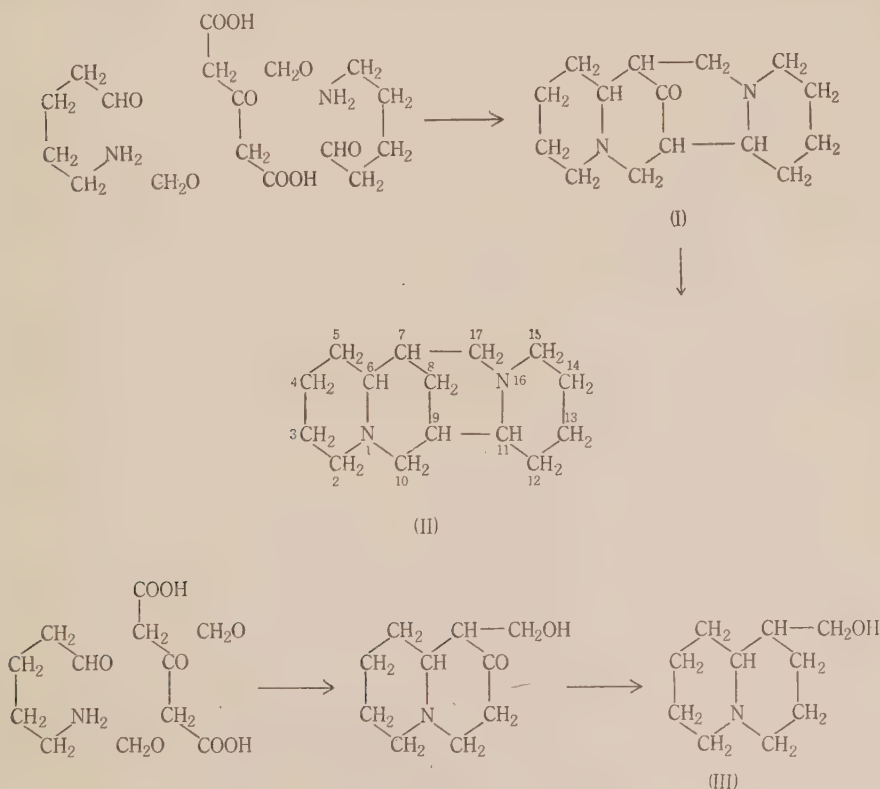
## Summary

The syntheses of sparteine and certain related derivatives of 1-azabicyclo-(4,4,0)decane are described and the stereochemistry of sparteine is discussed.

## I. INTRODUCTION

Since the biosynthesis of sparteine suggested by Robinson(1) in 1917 was based on the then prevailing erroneous conception of its structure, it was necessarily incorrect. Indeed it was not until 1936 that Clemo, Morgan, and Raper(2) demonstrated that its structure must be II and this has since been verified by several syntheses of the alkaloid(3-5).

However, a logical extension of Robinson's views on the biogenesis of alkaloids leads to the following formulation for the biosyntheses of sparteine (II) and lupinine (III).



\* A brief account of this work appeared in *Nature* **165**: 35 (1950).

<sup>†</sup> School of Chemistry, University of Sydney.



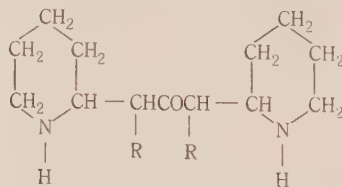
In the case of sparteine the last step cannot proceed through the stages: reduction of the carbonyl group to an alcohol, dehydration, and reduction, since the formation of the olefine would involve a contravention of Bredt's rule. However, in lupinine this may be the path followed. In any case the transformation of the carbonyl to a methylene group should be possible *in vivo*. The condensation reactions, being variations of the Mannich reaction, are obviously related to those proposed by Robinson for other alkaloids and moreover they clearly suggest a close relationship not only between lupinine, sparteine, and the alkaloids of the latter group but also between these alkaloids and those of the *isopelletierine* group.

Initially our attempts were directed at the synthesis of sparteine since it was anticipated that by-products would be formed in smaller amounts than in the synthesis of lupinine. When a dilute aqueous solution of 5-aminopentanal, formaldehyde, and acetonedicarboxylic acid at pH 7 was allowed to stand for several days at room temperature none of the expected product was formed. This result was probably due to the rapid condensation of the acid with formaldehyde followed by decarboxylation. To avoid these undesirable side reactions the condensation was therefore conducted in two stages. The aminoaldehyde was allowed to react with the acid in alkaline solution (cf. 6, 7) then the solution was acidified and formaldehyde added (cf. 8). The yields of *spartein-8-one*\* (I) obtained from the stoichiometric proportions are summarized below:

pH, first stage ..	7	11	13	13
pH, second stage ..	7	6	6	3
Yield (%) .. ..	0	0	18	30

Clemmensen reduction of I then gave *dl*-sparteine in high yield. The synthetic base formed a dipicrate, m.p. 205–207 °C. undepressed by admixture with authentic *dl*-sparteine dipicrate, m.p. 205–207 °C., and a monopicrate, m.p. 135 °C., also undepressed by authentic material of m.p. 135 °C.

The intermediate formed in the first stage of the synthesis of I was presumably the acid (IV, R=COOH), which in the second stage condensed with formaldehyde and then decarboxylated. This acid would probably be unstable

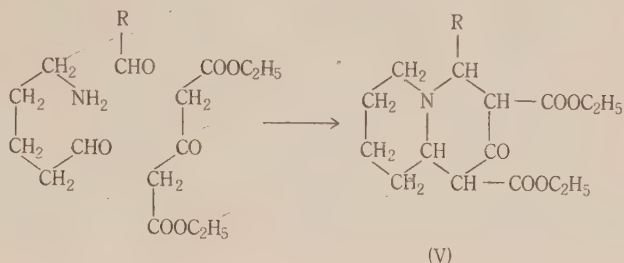


(IV)

and difficult to isolate but in support of this hypothesis, *1,3-bis(2'-piperidyl)-propanone* (IV, R=H) was isolated from the reaction between 5-aminopentanal and acetonedicarboxylic acid at pH 11. As expected this substance failed to condense with formaldehyde to I. (cf. 8).

\* This name is more systematic than 8-oxysparteine, cf. HENRY, T. A.—“The Plant Alkaloids.” 4th Ed. p. 138. (Churchill: London, 1949), or 8-oxosparteine (cf. 4) and is preferred.

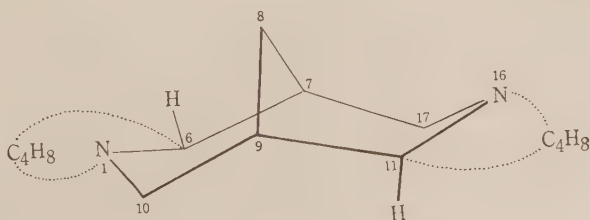
Syntheses of some related but simpler substances using 5-aminopentanal were also effected. These were prepared in one step at pH 3 (cf. 9) according to the equation



From formaldehyde, *diethyl 1-azabicyclo(4,4,0) decan-4-one-3,5-dicarboxylate* (V, R=H) was obtained in high yield, and this on hydrolysis and decarboxylation gave the known *1-azabicyclo(4,4,0) decan-4-one* (10). Similarly acetaldehyde furnished the *homologue* (V, R=CH<sub>3</sub>) and then *2-methyl-1-azabicyclo(4,4,0) decan-4-one*, from which by Clemmensen reduction, *2-methyl-1-azabicyclo(4,4,0) decane* was obtained. Benzaldehyde, however, afforded only a low yield of the *phenyl derivative* (V, R=C<sub>6</sub>H<sub>5</sub>) and all attempts to convert this to the parent base were unsuccessful.

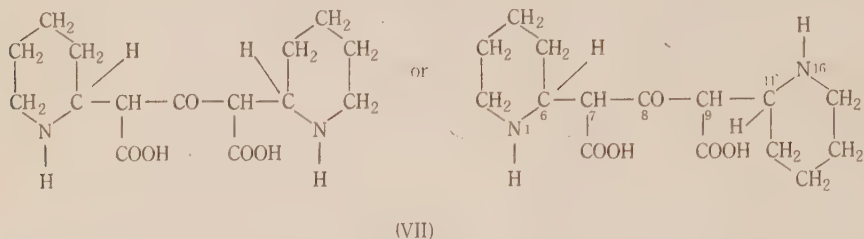
The synthesis of sparteine described above throws some light on the stereochemistry of the alkaloid. The initial condensation is essentially a Mannich type reaction and is therefore closely related to those reactions leading to the formation of certain 2,6-dialkyl-4-piperidones(11), lobelanine(6), and cuscohygrine(12). In the first and second of these cases it has been proved that the substances are *meso* forms, i.e. that the 2 and the 6 carbon atoms have mirror image configurations, and in cuscohygrine there is little doubt that an analogous situation obtains. Now in the syntheses of 1,3-bis(2'-piperidyl)propanone (IV, R=H) and its *N,N'*-dimethyl derivative(7) only one substance could be isolated in each instance. Hence if it be assumed that the same steric features are present in these two reactions as in those cited above it follows that these two substances *also* are *meso* forms and that in spartein-8-one and in sparteine, carbon atoms 6 and 11 have mirror image configurations.

From the evidence of scale models and the failure to isolate stereoisomeric forms of *bicyclo(3,3,1)nonane* and its derivatives it is accepted(13) that in such bicyclic systems the rings are *cis*. Therefore, in sparteine ring B is fused *cis* to ring C. This fact, together with the hypothesis that carbon atoms 6 and 11 have mirror image configurations, leads to the conclusion that sparteine has the configuration represented by VI.

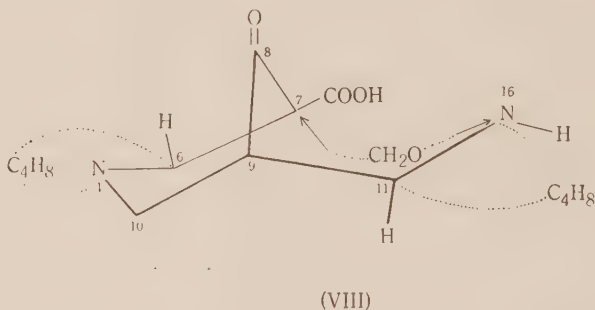


(VI)

The argument may perhaps be made clearer by the following considerations. The intermediate in the synthesis of I is VII in which carbon atoms 6 and 11 have mirror image configurations. Since the acid VII would exist partly in the enolic form the question of the configurations of carbon atoms 7 and 9 does not



arise. Now when a molecule of formaldehyde condenses with VII at carbon 9 and nitrogen 1 a molecule of carbon dioxide is evolved, carbon 9 becomes asymmetric, and ring *B* is formed. When the second molecule of formaldehyde condenses with carbon 7 and nitrogen 16 to form ring *C* it can do so in only *one* way and that is to make the fusion of ring *C* to *B* *cis*, as in the diagram (VIII) below. Hence only one racemate viz. VI is formed.



The same conclusion that sparteine is VI has been reached by Leonard and Beyler(4) on other grounds.

## II. EXPERIMENTAL

### (a) 4-Chlorobutanol

The crude aldehyde prepared by the method of Paul(14) using lead tetra-acetate was purified through its bisulphite compound. The overall yield from tetrahydrofurfuryl alcohol was 13–16%, b.p. 52 °C./14 mm.

Found: C, 45.0; H, 6.7; Cl, 33.4%.

Calculated for  $C_4H_7OCl$ : C, 45.1; H, 6.6; Cl, 33.3%.

The residual portion which had not reacted with bisulphite had b.p. 52 °C./14 mm. and  $n_D^{20}$  1.4556, and was obtained in 10–20% overall yield.

Found: C, 49.6; H, 7.7; Cl, 29.7%.

Calculated for  $C_5H_9OCl$ : C, 49.8; H, 7.5; Cl, 29.4%.

It is probably tetrahydrofurfuryl chloride (lit. b.p. 47–48 °C./15 mm., and  $n_D^{20}$  1.4560).

For the preparation of its acetal it is not necessary to purify the 4-chlorobutanol.

(b) *Sparteïn-8-one (I)*

A solution of 5-amino-1,1-diethoxypentane (12 g.; 2 mol.) in hydrochloric acid (10%), just acid to Congo red, was warmed on the water-bath for 8 minutes, then quickly cooled, neutralized, and added to a neutralized solution of acetonedicarboxylic acid (5 g.; 1 mol.). Water and sodium hydroxide solution were added to bring the volume to 200 ml. and the pH to the required level. After standing at room temperature for 3 days the pH was adjusted to the required value by the addition of cold hydrochloric acid, formaldehyde (5.2 g. of 40%; 2 mol.) added, followed by a concentrated buffer solution, the final volume being 350 ml. After a further 3 days at room temperature the mixture was saturated with potassium carbonate and thoroughly extracted with ether, much amorphous material remaining undissolved. The product isolated by distillation under reduced pressure had b.p. 160 °C./1 mm. and solidified on standing. The yields are given in the introduction. The substance crystallized from acetone or light petroleum (b.p. <40 °C.) in colourless plates, m.p. 135 °C. which contained one molecule of water of crystallization.

Found: C, 67.4; H, 10.0; N, 10.5%.

Calculated for  $C_{15}H_{24}ON_2H_2O$ : C, 67.6; H, 9.9; N, 10.5%.

The water of crystallization was not completely lost by drying at 100 °C./1 mm., and at higher temperatures the substance began to sublime, thus rendering an accurate determination impossible. The partially dried material readily took up moisture on short exposure to air and regained its original weight.

(c) *Sparteïne (II)*

Sparteïn-8-one (0.5 g.), concentrated hydrochloric acid (10 ml.), and amalgamated zinc (5 g.) were refluxed for 20 hours, additional amounts of hydrochloric acid (0.5 ml.) being added every 3 hours. After cooling, the supernatant liquid was decanted, made strongly alkaline, and exhausted with ether. The product (0.4 g.; 80%) was a colourless oil, b.p. 160 °C./10 mm.

The *dipicrate* crystallized from alcohol as yellow needles, m.p. 205–207 °C., undepressed by admixture with authentic *dl*-sparteïne dipicrate, m.p. 205–207 °C.

Found: C, 46.7; H, 4.8; N, 16.3%.

Calculated for  $C_{27}H_{32}O_{14}N_8$ : C, 46.8; H, 4.7; N, 16.2%.

The *monopicate* separated from alcohol in yellow needles, m.p. 135 °C., undepressed by admixture with authentic *dl*-sparteïne monopicate, m.p. 135 °C.

Found: N, 14.9%.

Calculated for  $C_{21}H_{29}O_7N_5$ : N, 15.1%.

(d) *1,3-Bis(2'-Piperidyl)Propanone (IV, R=H)*

5-Aminopentanal was condensed with acetonedicarboxylic acid at pH 11 for 3 days as above. The product (35% yield), isolated by ether extraction after saturation with potassium carbonate, was a very pale yellow viscous oil, b.p. 120 °C./0.1 mm., which rapidly darkened in air.

The *dipicrolonate*, precipitated from alcohol, was very sparingly soluble in most solvents but crystallized from alcohol-nitrobenzene in stout yellow needles, m.p. 231–232 °C. (decomp.).

Found: C, 52.5; H, 5.2; N, 18.4%.

Calculated for  $C_{33}H_{40}O_{11}N_{10}$ : C, 52.6; H, 5.4; N, 18.6%.

(e) *Diethyl 1-Azabicyclo(4,4,0)Decan-4-One-3,5-Dicarboxylate (V, R=H)*

To a solution of 5-amino-1,1-diethoxypentane (2 g.; 1 mol.) in alcohol (8 ml.), made just acid to methyl orange with hydrochloric acid (10%), was added ethyl acetonedicarboxylate (2.3 g.; 1 mol.) and formaldehyde (1 g. of 35%; 1 mol.). The homogeneous reaction mixture was allowed to stand at room temperature for 2 days, then made alkaline to phenolphthalein with potassium carbonate, and extracted with ether. The ethereal solution was washed three times with water to remove alcohol and then extracted with hydrochloric acid (3%). The acid solution was basified with potassium carbonate and the liberated base taken up in ether. After



drying and removing the ether a viscous pale yellow liquid (3.1 g.; 90% yield) remained. It could not be distilled without decomposition, did not yield a crystalline picrate or picrolonate, and darkened on exposure to air. However, the derivative described below was readily obtained.

(f) *1-Azabicyclo(4,4,0)Decan-4-One*

To a solution of V ( $R=H$ ) (3 g.) in hydrochloric acid (20 ml. of 5N) were added a few fragments of amalgamated zinc and the mixture refluxed for 8 hours. After cooling, the mixture was basified with potassium carbonate, extracted with ether, and the base eventually isolated by distillation. The product (0.5 g.; 30% yield) was a colourless oil, b.p. 70 °C./1 mm. which darkened on exposure to air.

The *picrate* crystallized from acetone or alcohol in old-gold prisms, m.p. 208 °C. (lit. 211 °C.).

Found: N, 14.6%.

Calculated for  $C_{15}H_{18}O_8N_4$ : N, 14.7%.

(g) *Diethyl 2-Methyl-1-Azabicyclo(4,4,0)Decan-4-One-3,5-Dicarboxylate* ( $V$ ,  $R=CH_3$ )

The condensation was carried out as above using acetaldehyde in place of formaldehyde and the crude product (95% yield) isolated in the same way. It was purified by distillation under reduced pressure, forming a colourless oil, b.p. 140 °C./0.001 mm.

The *picrolonate* crystallized from alcohol in yellow blades, m.p. 160 °C.

Found: N, 12.2%.

Calculated for  $C_{26}H_{33}O_{10}N_5$ : N, 12.2%.

(h) *2-Methyl-1-Azabicyclo(4,4,0)Decan-4-One*

Hydrolysis and decarboxylation of V ( $R=CH_3$ ) (14.5 g.) was effected as in the case of V ( $R=H$ ). The product (4.5 g.; 60% yield) was a colourless mobile oil, b.p. 85 °C./1 mm., which darkened in air.

Found: N, 8.3%.

Calculated for  $C_{10}H_{17}ON$ : N, 8.4%.

The *picrate* crystallized from acetone in yellow needles, m.p. 189 °C.

Found: N, 14.0%.

Calculated for  $C_{16}H_{20}O_8N_4$ : N, 14.1%.

(i) *2-Methyl-1-Azabicyclo(4,4,0)Decane*

The above ketone (1.7 g.), hydrochloric acid (10 ml. of 5N), and amalgamated zinc (5 g.) were refluxed for 20 hours, with the addition of hydrochloric acid (1 ml.) every 3 hours. The base (1.3 g.; 80% yield) isolated in the usual manner was a colourless oil, b.p. 75 °C./8 mm., darkening on exposure to air.

The *picrate* separated from alcohol in yellow plates, m.p. 165 °C.

Found: N, 14.8%.

Calculated for  $C_{16}H_{22}O_7N_4$ : N, 14.7%.

The *hydrobromide* crystallized from acetone in colourless needles which began to darken at 290 °C. and finally melted at 315 °C.

Found: N, 6.0%.

Calculated for  $C_{10}H_{20}NBr$ : N, 6.0%.

(j) *Diethyl 2-Phenyl-1-Azabicyclo(4,4,0)Decan-4-One-3,5-Dicarboxylate* ( $V$ ,  $R=C_6H_5$ )

The condensation of 5-aminopentanal (5.2 g.; 1 mol.), ethyl acetonedicarboxylate (6.2 g.; 1 mol.), and benzaldehyde (3.2 g.; 1 mol.) was carried out in the usual way. The crude product, a pale yellow viscous oil, obtained in low yield (3 g.; 25%) could not be distilled.

The *picrate* separated from alcohol in yellow prisms, m.p. 172 °C.

Found: N, 9.2%.

Calculated for  $C_{27}H_{30}O_{12}N_4$ : N, 9.3%.

## III. ACKNOWLEDGMENTS

The authors are grateful to Miss J. Fildes and Miss B. Naylor for the analyses, to the Commonwealth Research Grant Committee of the University of Sydney for a scholarship awarded to one of them (E.F.L.J.A.), and to the Quaker Oats Company through their distributors, Swift and Company Pty. Ltd., for a generous gift of tetrahydrofurfuryl alcohol.

## IV. REFERENCES

- (1) ROBINSON, R.—*J. Chem. Soc.* **111** : 876 (1917).
- (2) CLEMO, G. R., MORGAN, W. MCG., and RAPER, R.—*J. Chem. Soc.* **1936** : 1025 (1936).
- (3) CLEMO, G. R., RAPER, R., and SHORT, W. S.—*J. Chem. Soc.* **1949** : 663 (1949).
- (4) LEONARD, N. J., and BEYLER, R. E.—*J. Amer. Chem. Soc.* **72** : 1316 (1950).
- (5) SORM, F., and KEIL, B.—*Chem. Abstr.* **43** : 3828 (1949).
- (6) SCHÖPF, C., and LEHMANN, G.—*Liebigs Ann.* **518** : 1 (1935).
- (7) ANET, E. F. L. J., HUGHES, G. K., and RITCHIE, E.—*Aust. J. Sci. Res. A* **3** : 336 (1950).
- (8) ANET, E. F. L. J., HUGHES, G. K., MARMION, DIANA, and RITCHIE, E.—*Aust. J. Sci. Res. A* **3** : 330 (1950).
- (9) LIONS, F., and WILLISON, A. M.—*J. Roy. Soc. N.S.W.* **73** : 240 (1939).
- (10) CLEMO, G. R., METCALFE, T. P., and RAPER, R.—*J. Chem. Soc.* **1936** : 1429 (1936).
- (11) MANNICH, C.—*Arch. Pharm. Berl.* **272** : 323 (1934).
- (12) ANET, E. F. L. J., HUGHES, G. K., and RITCHIE, E.—*Aust. J. Sci. Res. A* **2** : 616 (1949).
- (13) GILMAN, H.—“Organic Chemistry.” 2nd Ed. Vol. 1, p. 486. (Wiley & Sons: New York, 1944.)
- (14) PAUL, R.—*C.R. Acad. Sci. Paris*, **211** : 645 (1940); **215** : 303 (1942).

# AROMATIC ALDEHYDES FROM THE OXIDATION OF SOME AUSTRALIAN WOODS AND THEIR CHROMATOGRAPHIC SEPARATION

By D. E. BLAND,\* G. HO,† and W. E. COHEN\*

[Manuscript received August 8, 1950]

## Summary

The mixtures of aromatic aldehydes produced from 12 species of eucalypt woods and eight Australian gymnosperm woods by oxidation with nitrobenzene and alkali were separated by sublimation and the individual aldehydes identified. A paper partition chromatographic method for the separation of the aldehydes was used to confirm the results of the sublimation separation and to identify traces of aldehydes in some of the mixtures. The eucalypt woods all yielded vanillin, syringaldehyde, and traces of *p*-hydroxybenzaldehyde. The gymnosperm woods yielded mainly vanillin but traces of syringaldehyde were produced from some and traces of *p*-hydroxybenzaldehyde from most.

## I. INTRODUCTION

The production of an aromatic aldehyde from lignin was first noted by Grafe(1) who found that, when the solid residue from sulphite waste liquor was heated with lime in a sealed tube, vanillin could be recovered from the reaction mixture. Since then many modifications of Grafe's procedure have been proposed, most of them being aimed at the commercial production of vanillin as a by-product of the sulphite pulping of gymnosperm woods. These processes in general consist of heating the concentrated sulphite waste liquor with sodium hydroxide(1). Freudenberg, Lautsch, and Engler(2) found that, by the addition of nitrobenzene to the alkaline liquor, a greatly improved yield of vanillin could be obtained and also that similar treatment of wood gave a high yield of vanillin. Creighton, McCarthy, and Hibbert(3) and Creighton, Gibbs, and Hibbert(4) made use of this oxidation technique and confirmed the results of Freudenberg. They found, however, that syringaldehyde was produced in greater quantities than vanillin from the wood of angiosperms. Creighton and Hibbert(5) applied the same oxidation to cornstalks and found, among the products, syringaldehyde, vanillin, and *p*-hydroxybenzaldehyde. They suggested that the presence of this last compound might be a distinguishing characteristic of monocotyledons.

In the present investigation 12 species of eucalypt (dicotyledon) woods were oxidized by alkaline nitrobenzene and the resulting aldehydes separated and identified. A paper partition chromatographic method, for the separation of the aldehydes, already published by Bland(6) was used to confirm the results obtained by sublimation separation and to test for aldehydes present in small

\* Division of Forest Products, C.S.I.R.O., Melbourne.

† Seconded from National Bureau of Forest Research, Nanking, China.

quantities only. The "Chromatopile" of Mitchell and Haskins(7) was used to isolate an aldehyde present in small amount.

Eight gymnosperm woods were similarly oxidized and the aldehydes analysed chromatographically. The aldehydes from three of these woods were separated by sublimation.

## II. EXPERIMENTAL

All melting points given are corrected.

### (a) Oxidation of Woods

Air-dry sawdust (5 g.) of the species under investigation, nitrobenzene (3.0 ml.), and 2N sodium hydroxide (100 ml.) were added to a steel bomb of 200 ml. capacity. The bomb was sealed, shaken vigorously for a few seconds, and agitated in an oil-bath for 3 hours at 150 °C.

### (b) Isolation of Mixed Aldehydes

The reaction mixture from the oxidation was filtered and the residue washed twice with a small quantity of distilled water. The filtrate was extracted with benzene to remove excess nitrobenzene and its reduction products, and the aqueous solution made acid with a 10% excess hydrochloric acid and allowed to stand for 24 hours. It was then filtered to remove the precipitate which had formed; this amounted to 13 to 25% of the original wood.

The clear filtrate was extracted with benzene and the combined benzene washings concentrated to a convenient volume (c. 150 ml.). This concentrate was extracted with 20% sodium bisulphite solution until no more aldehyde could be detected in the washings by acidifying, boiling off the sulphur dioxide, and testing with 2,4-dinitrophenylhydrazine. The combined bisulphite washings were acidified with hydrochloric acid and the sulphur dioxide removed by heating in a stream of carbon dioxide. They were then extracted with benzene until free from aldehyde. The benzene extracts were concentrated to small volume, transferred quantitatively to a Pyrex evaporating dish, and taken to dryness on a water-bath and dried in an oven at 105 °C. for 5 minutes. This left the aldehydes as an oil which set solid on cooling and became crystalline after standing a few days.

### (c) Separation of Aldehydes by Sublimation

A weighed sample of the aldehyde mixture (c. 0.2 g.) was placed in a tube fitted with a cold finger. The tube was placed in a bath at 61 °C. and kept evacuated to about 1.5 mm. for 8 hours, by which time the vanillin had sublimed onto the cold finger. It was removed by breaking the vacuum and scraping from the cold finger into a small tared Pyrex basin. The last trace of vanillin was washed into the basin with a few drops of hot alcohol. The alcohol was evaporated on the water-bath, the dish wiped dry, placed in an oven at 105 °C. for 5 minutes, and allowed to cool in a desiccator. The recovered vanillin was weighed as such. Syringaldehyde was collected in the same way by continuing the sublimation at 100 °C. for about 12 hours.

### (d) Identification of the Sublimates

The sublimate obtained at 61 °C. (m.p. 75–77 °C.), when recrystallized from light petroleum, b.p. 100–110 °C., yielded colourless needles, m.p. 81 °C.; mixed m.p. with vanillin, 81 °C.

Found:  $\text{CH}_3\text{O}$ , 20.1%.

Calculated for  $\text{C}_8\text{H}_8\text{O}_3$ :  $\text{CH}_3\text{O}$ , 20.4%.

Absorption maximum in ultraviolet was at 355  $\mu$  (lit. 353  $\mu$ )(8). It yielded a 2,4-dinitrophenylhydrazone, m.p. 264 °C., which gave crystals, m.p. 272 °C. (lit. 271 °C.) from glacial acetic acid.

The sublimate obtained at 100 °C., m.p. 111–112 °C., when recrystallized five times from water, gave pale yellow needles, m.p. 113 °C. (lit. syringaldehyde, 111–114 °C.).

Found:  $\text{CH}_3\text{O}$ , 34.2%.



Calculated for  $C_9H_{10}O_4$ :  $CH_3O$ , 34.1%.

Absorption maximum in ultraviolet was at 370  $m\mu$  (lit. 370  $m\mu$ )(8). Its 2,4-dinitrophenylhydrazone gave crystals, m.p. 235 °C. (lit. 235 °C.) from glacial acetic acid.

(e) *Paper Partition Chromatographic Separation of the Aldehydes*

This was carried out using the method already published elsewhere(6). Three different solvents were employed: Light petroleum, b.p. 100–110 °C.; a mixture of "Shell solvent X222", b.p. 40–100 °C., and benzene (8 : 2); and "Shell solvent X6", b.p. 105–115 °C. As an alternative to spraying the sheets the position of the aldehyde spots was disclosed by drawing the dried sheet across the surface of a 0.1% solution of 2,4-dinitrophenylhydrazine in 2N hydrochloric acid contained in a photographic developing dish. Analyses were carried out on the isolated aldehydes by spotting the paper with a 2% solution of the aldehydes in ethanol and also on the crude solution obtained by concentrating the reaction mixture after acidification and removal of the precipitate.

(f) *Isolation of a Third Aldehyde*

Air dry *E. corymbosa* Sm. sawdust (500 g.) was heated in an autoclave with nitrobenzene (300 ml.) and 2N sodium hydroxide (10 l.) for 3 hours at 150 °C. After cooling, the charge was acidified with 10% excess hydrochloric acid and allowed to stand overnight. It was filtered to remove the bulky residual pulp and precipitate. A paper chromatograph of the filtrate showed the presence of three aldehydes. It was extracted with chloroform and the progress of removal of the aldehydes followed on a paper chromatograph. After 5 washes the vanillin and syringaldehyde spots had almost disappeared and the third aldehyde spot was still present. After extraction with ether the chromatograph showed that the third aldehyde had passed completely into the ether.

The ether solution was concentrated and the concentrate soaked up by a sheet of bleached soda eucalypt blotter (10 by 10 cm.) which, after drying, became the loaded sheet of a "Chromatopile". The capillary ascent technique was used simply by standing the pile of blotter sheets on the bottom of a stainless steel chromatography tank with a weight on top of the pile. The solvent ("Shell solvent X222": benzene (1 : 1)) was poured into the tank where it covered the bottom to a depth of about 2 cm. An open vessel containing water equilibrated with the solvent was placed on the bottom of the tank.

After standing overnight the pile was taken from the tank, spread out, and allowed to dry. A half-inch strip was cut from the edge of each of the sheets and dipped in a 0.1% solution of 2,4-dinitrophenylhydrazine in 2N hydrochloric acid. The sheets whose strip gave a colour with this reagent were torn up and extracted three times with boiling water. The water extracts were cooled, filtered, and extracted with ether. The ether solution was concentrated and the concentrate allowed to evaporate in air. This gave 0.10 g. of pinkish crystals.

When recrystallized from benzene this substance melted at 114 °C.; with *p*-hydroxybenzaldehyde it gave a mixed m.p. of 115 °C. (lit. 116 °C.) and with syringaldehyde, 79 °C. Absorption maximum in ultraviolet was at 335  $m\mu$  (lit. 336  $m\mu$ )(8). It gave a 2,4-dinitrophenylhydrazone which when recrystallized from glacial acetic acid melted at 279 °C. (decomp.), and gave a mixed m.p. with the 2,4-dinitrophenylhydrazone of *p*-hydroxybenzaldehyde of 280 °C. (decomp.) (lit. 280 °C. decomp.).

### III. RESULTS AND DISCUSSION

The results of the separation by sublimation of the aldehydes from 12 species of eucalypt woods are shown in Table 1. The yields are similar to those obtained by Creighton, McCarthy, and Hibbert(3) and Creighton, Gibbs, and Hibbert(4) from North American hardwoods, except that the yield from *Eucalyptus marginata* Sm. is lower than that from the other eucalypts and nearer to that obtained by those authors from spruce.

TABLE I  
ALDEHYDES FROM EUCALYPT WOODS AND THEIR SEPARATION BY SUBLIMATION

Wood	Oven-Dry Weight (g.)	Lignin (%)	Aldehydes		Results of Sublimation Separation of Aldehydes				Ratio (S/V)†
			Recovered (g.)	Lignin (%)	Vanillin (%)	Syring-aldehyde (%)	Residue (%)	Loss (%)	
<i>E. regnans</i> F.v.M.—young	14.4	22.3	1.35	42.0	13.1	67.5	13.1	6.3	4.3
					15.8	70.8	8.5	4.9	3.7
<i>E. regnans</i> F.v.M.—mature	18.3	23.9	1.44	32.9	13.9	73.1	4.6	8.4	4.4
					15.6	74.2	4.5	5.7	4.0
					14.7	73.6	4.8	6.9	4.2
<i>E. gigantea</i> Hook. f.	17.6	20.1	1.44	40.7	16.5	77.6	4.7	1.2	3.9
<i>E. obliqua</i> L'Herit.	18.2	16.4	1.67	55.9	16.6	78.4	5.0	0.0	3.9
<i>E. marginata</i> Sm.	16.6	34.8	1.24	21.5	27.3	49.8	3.6	19.3	1.5
					29.2	53.8	4.4	12.6	1.5
					34.4	54.5	8.0	3.1	1.3
<i>E. patens</i> Benth.	18.1	26.5*	1.10	22.9	34.4	54.5	8.0	3.1	1.3
<i>E. pilularis</i> Sm.	17.7	22.1*	1.42	36.2	23.1	63.5	13.2	0.2	2.3
	18.0	22.1*	1.27	31.8	23.1	63.5	8.8	4.6	2.3
<i>E. maculata</i> Hook.	18.1	18.9*	1.08	31.6	12.9	79.8	6.6	0.7	5.2
<i>E. calophylla</i> R.Br.	17.5	22.8*	1.00	25.1	20.5	68.5	10.1	0.9	2.8
<i>E. corymbosa</i> Sm.	18.3	20.8*	1.19	31.3	22.5	66.1	8.8	2.6	2.5
<i>E. resinifera</i> Sm.	17.8	24.2*	1.48	34.4	18.9	65.0	12.2	3.9	2.9
<i>E. diversicolor</i> F.v.M.	18.2	18.0*	1.51	46.1	12.7	79.3	8.6	+0.6	5.2
<i>E. saligna</i> Sm.	17.9	25.4*	1.40	30.8	20.1	68.1	8.3	3.5	2.8

\* Wood pre-extracted with alkali.

† Moles of syringaldehyde per mole of vanillin.

The properties of the two principal products of the sublimation identify them as consisting mainly of vanillin and syringaldehyde although application of the chromatographic separation to the sublimates disclosed a small amount of vanillin in the syringaldehyde sublimate and *vice versa*. The results given in Table 1 have been calculated in terms of moles of syringaldehyde per mole of vanillin. It may be seen by reference to this table that the ratio found varied from 1.3 to 5.2 for the specimens of different species tested.

Only vanillin could be identified by subliming the aldehydes from the gymnosperms *Pinus radiata* D. Don., *Araucaria cunninghamii* Ait., and *Phyllocladus rhomboidalis* L. C. Rich.

The results of the paper partition chromatographic separation of the aldehydes by means of light petroleum/water gave some indication of the existence of a third substance in the aldehyde mixture because, on treating the paper with 2,4-dinitrophenylhydrazine, a spot developed on the site of the original spot. That this was due to the presence of a third aldehyde was confirmed by the results with the other two solvents, both of which gave a measurable  $R_F$  value with it. Pure *p*-hydroxybenzaldehyde gave an  $R_F$  value identical with the third aldehyde in all three solvent systems. The results are shown in Table 2.

TABLE 2  
 $R_F$  VALUES OF ALDEHYDES IN DIFFERENT SOLVENT SYSTEMS

Solvent	Vanillin	Syringaldehyde	<i>p</i> -Hydroxybenzaldehyde
Light petroleum (b.p. 100–110 °C.) ..	0.32	0.12	0.00
"Shell solvent X6" .. .. .	0.85	0.55	0.12
"Shell solvent X222" + benzene (8 : 2)	0.54	0.25	0.04

As the third aldehyde had a low  $R_F$  value corresponding to a high partition coefficient (concentration in water phase/concentration in organic phase) the procedure of isolating the aldehydes by extraction with benzene may have resulted in some loss of the third aldehyde. This seems to be borne out by the fact that in some cases it was detectable in the aqueous concentrate but not in the isolated aldehydes. The results obtained on 12 eucalypts are shown in Table 3; it can be seen by reference to this table that the third aldehyde was detected as a product from all 12 species. In some cases the third substance was not detected in the isolated aldehyde mixture but was detectable in the aqueous concentrate; e.g. in the case of the three "ash" eucalypts the third aldehyde showed only as a faint spot (queried in Table 3). Treatment of vanillin and of syringaldehyde under the same conditions as the original oxidation did not produce any of the third aldehyde.

Chromatographic analysis of the crude concentrates prepared from seven native and one exotic (*P. radiata*) gymnosperm woods showed detectable amounts

TABLE 3  
RESULTS OF PAPER PARTITION CHROMATOGRAPHIC SEPARATION OF ALDEHYDES FROM VARIOUS  
EUCALYPT WOODS

Wood	Aldehyde Sample	Vanillin	Syring- aldehyde	Third Aldehyde
<i>E. patens</i> Benth. .. ..	Isolated ..	+	+	—
	Crude ..	+	+	+
<i>E. pilularis</i> Sm. .. ..	Crude ..	+	+	+
<i>E. maculata</i> Hook. .. ..	Crude ..	+	+	+
<i>E. calophylla</i> R.Br. .. ..	Crude ..	+	+	+
<i>E. corymbosa</i> Sm. .. ..	Isolated ..	+	+	+
	Crude ..	+	+	+
<i>E. marginata</i> Sm. .. ..	Isolated ..	+	+	—
	Crude ..	+	+	+
<i>E. resinifera</i> Sm. .. ..	Isolated ..	+	+	?
	Crude ..	+	+	+
<i>E. saligna</i> Sm. .. ..	Isolated ..	+	+	+
	Crude ..	+	+	+
<i>E. diversicolor</i> F. v. M. ..	Crude ..	+	+	+
<i>E. obliqua</i> L'Herit. .. ..	Isolated ..	+	+	—
	Crude ..	+	+	?
<i>E. regnans</i> F. v. M. .. ..	Isolated ..	+	+	—
	Crude ..	+	+	?
<i>E. gigantea</i> Hook. f. .. ..	Isolated ..	+	+	—
	Crude ..	+	+	?

of the third aldehyde in five of these, traces in two, and no detectable amount in one of the native woods. These results, shown in Table 4, are notable in that they disclose the presence of syringaldehyde in the aldehydes from three of the gymnosperm woods.

TABLE 4  
RESULTS OF PAPER PARTITION CHROMATOGRAPHIC SEPARATION OF ALDEHYDES FROM GYMNOSPERM  
WOODS

Wood	Aldehyde Sample	Vanillin	Syring- aldehyde	Third Aldehyde
<i>Callitris glauca</i> R.Br. ..	Crude	+	?	?
<i>Agathis palmerstoni</i> F.v.M. . .	„	+	—	+
<i>Podocarpus elata</i> R.Br. ..	„	+	+	+
<i>Athrotaxis selaginoides</i> D.Don.	„	+	?	+
<i>Phyllocladus rhomboidalis</i> L.C.Rich .. ..	„	+	—	?
<i>Dacrydium franklinii</i> Hook f.	„	+	—	—
<i>Araucaria cunninghamii</i> Ait.	„	+	—	+
<i>Pinus radiata</i> D.Don. ..	„	+	—	+

9  
1

16

93



## IV. CONCLUSIONS

The relative amounts of the different aromatic aldehydes produced on nitrobenzene oxidation of eucalypt woods varied greatly between specimens of different species. There appears to be at least as much variation as that already found among North American angiosperms. Eucalypt lignin cannot therefore be regarded as a chemical entity. The statement that the woods of the angiosperms yield vanillin and syringaldehyde, whereas the woods of gymnosperms yield only vanillin, needs qualification to allow for traces of syringaldehyde from some of the gymnosperms and for traces of *p*-hydroxybenzaldehyde from both angiosperms and gymnosperms. The presence of *p*-hydroxybenzaldehyde is not a distinguishing characteristic of monocotyledons if its presence in small quantities is taken into consideration.

## V. ACKNOWLEDGMENTS

The authors wish to thank Miss J. Harvey and Miss A. McCance for their assistance.

## VI. REFERENCES

- (1) WISE, L. E.—“Wood Chemistry.” pp. 321-5. (Reinhold: New York, 1944.)
- (2) FREUDENBERG, K., LAUTSCH, W., and ENGLER, K.—*Ber. dtsh. chem. Ges.* **73** B: 167 (1940).
- (3) CREIGHTON, R. H. J., MCCARTHY, J. L., and HIBBERT, H.—*J. Amer. Chem. Soc.* **63**: 3052 (1941).
- (4) CREIGHTON, R. H. J., GIBBS, R. D., and HIBBERT, H.—*J. Amer. Chem. Soc.* **66**: 32 (1944).
- (5) CREIGHTON, R. H. J., and HIBBERT, H.—*J. Amer. Chem. Soc.* **66**: 37 (1944).
- (6) BLAND, D. E.—*Nature* **164**: 1093 (1949).
- (7) MITCHELL, H. K., and HASKINS, F. A.—*Science* **110**: 278 (1949).
- (8) LEMON, H. W.—*J. Amer. Chem. Soc.* **69**: 2998 (1947).

# INDEX

	PAGE		PAGE
Absorption Spectra in Linear Polyenes .. .. .	109	Angyal, S. J., and Jenkin, S. R.— Sulphonamides. I. "Marfanil" and its <i>o</i> - and <i>m</i> -Isomers..	461
Absorption Spectra, Ultraviolet, of the Acridone Alkaloids	593, 615	<i>Anthotroche</i> Species, Ursolic Acid from .. .. .	516
Acetone and Ammonia, Re- action between .. .. .	450	Antimony, A.C. Polarography of .. .. .	567
Acridone Alkaloids ..	593, 615	Aromatic Aldehydes from the Oxidation of Some Australian Woods .. .. .	642
Acridone Nucleus, Compounds containing the .. .. .	593	Aromatic Compounds and Hydrocarbons, The Iodination of .. .. .	587
Acronycidine, The Structure and Reactions of .. .. .	155	Atmospheric Circulation, The Dynamics of .. .. .	1
Adsorption Equilibria ..	290		
Aldehydes, Aromatic, from the Oxidation of Some Australian Woods .. .. .	642	Bayliss, N. S.— Conjugated Compounds. II. Simple Potential Energy Functions, Absorption Spectra, and Ionization in Linear Polyenes .. .. .	109
Alkaloids, Acridone, Ultraviolet Absorption Spectra of the	593, 615	Bismuth, A.C. Polarography of	567
Alkaloids of the Australian Rutaceae .. .. .	155	Bland, D. E., Ho, G., and Cohen, W. E.— Aromatic Aldehydes from the Oxidation of Some Aus- tralian Woods and their Chromatographic Separa- tion .. .. .	642
Alkaloids of <i>Pleogyne cun- ninghamii</i> .. .. .	346	Bolton, J. G., and Westfold, K. C.— Galactic Radiation at Radio Frequencies. I. 100 Mc/s. Survey ..	19
Allen, C. W., and Gum, C. S.— Survey of Galactic Radio- Noise at 200 Mc/s. .. .. .	224	III. Galactic Structure	251
Alloys containing Two Phases, Deformation and Recrystal- lization of .. .. .	72	Bottomley, W., and White, D. E.— The Chemistry of Western Australian Plants. III. Ursolic Acid from <i>Antho- troche</i> Species .. .. .	516
Alloys, The Superstructure in the $\alpha$ Phase of Silver-Mag- nesium .. .. .	284	Bowen, E. G.— The Formation of Rain by Coalescence .. .. .	193
Ammonia and Acetone, Re- action between .. .. .	450		
Anet, E. F. L. J., Hughes, G. K., and Ritchie, E.— A Synthesis of <i>iso</i> Pelletierine and Methylisopelletierine..	336		
A Synthesis of Sparteine and Some Related Substances	635		
Anet, E. F. L. J., Hughes, G. K., Marmion, Diana, and Ritchie, E.— Some Bridged Derivatives of 4-Piperidone .. .. .	330		
Anet, F. A. L., Hughes, G. K., and Ritchie, E.— The Alkaloids of <i>Pleogyne cunninghamii</i> .. .. .	346		

	PAGE		PAGE
Breyer, B., Gutman, F., and Hacobian, S.—		Cooke, R. G., and Segal, W.—	
Polarography with Alternating Currents.		Colouring Matters of Australian Plants. I. The Structure of Droserone ..	628
I. Outline of Theory, Apparatus, and Technique	558	Cooke, R. G., and Somers, T. C. Dunnione and Related Naphthoquinones.	
II. A.C. Polarography of Cadmium, Zinc, Lead, Thallium, Indium, Bismuth, and Antimony ..	567	I. The <i>iso</i> Dunnione Series	466
Brown, I., and Ewald, A. H.—		II. Rearrangement of Hydroxynaphthoquinones to Indenone Carboxylic Acids. <i>allo</i> Dunnione	487
Liquid-Vapour Equilibria.		Correction .. ..	350
I. The Systems Carbon Tetrachloride - <i>cyclo</i> Hexane and Water-Acetic Acid ..	306	Corrigenda .. ..	172
Brown, R. D.—		Cosmic Rays at Sea-Level, Remarks on the Latitude Effect of .. ..	183
Asymptotic Expressions for the Energies of Certain Long Molecules .. ..	428	Crystal Structure of Indium Monobromide .. ..	581
Brown, R. D., and Lahey, F. N. The Ultraviolet Absorption Spectra of the Acridone Alkaloids.		Deacon, E. L.—	
I. Compounds containing the Acridone Nucleus ..	593	Radiative Heat Transfer in the Air near the Ground ..	274
II. Compounds related to 4-Quinolone .. ..	615	6 - ( $\gamma\gamma$ - Dimethylallyl) - 7 - Methoxycoumarin .. ..	342
Butadiene .. ..	290	Diacetonamines .. ..	450
2,3-Butanediol .. ..	290	Donald, Heather B.— <i>See</i> Kan-nuluik, W. G. .. ..	417
Cadmium, A.C. Polarography of Carbon Tetrachloride - <i>cyclo</i> - Hexane and Water-Acetic Acid Systems .. ..	306	Droserone, The Structure of ..	628
Clarebrough, L. M.—		<i>iso</i> Dunnionol .. ..	481
Deformation and Recrystallization of Alloys containing Two Phases ..	72	Dunnione .. ..	466, 481, 487
Clarebrough, L. M., and Nicholas, J. F.—		Energy Functions in Linear Polyenes .. ..	109
The Superstructure in the $\alpha$ Phase of Silver-Magnesium Alloys .. ..	284	Equilibria, Adsorption ..	290
Cohen, W. E.—		Equilibria, Liquid-Vapour ..	306
<i>See</i> Bland, D. E., and Ho, G.	642	Esters, A Colorimetric Method for the Determination of ..	128
<i>See</i> Foster, D. H., and Schwerin, Gerda .. ..	504	Ethylene to Ethylene Oxide, Oxidation of .. ..	433
Conductivity, Heat, in Gases ..	417	<i>Eucalyptus regnans</i> .. ..	504
Conjugated Compounds. II ..	109	Ewald, A. H.— <i>See</i> Brown, I. . .	306
Cooke, R. G.—		Ewing, Jean, Hughes, G. K., and Ritchie, E.—	
Dunnione and Related Naphthoquinones. II.		The Chemical Constituents of Australian <i>Xanthoxylum</i> Species.	
Synthesis of <i>iso</i> Dunnionol and <i>dl</i> -Dunnione .. ..	481	I. Suberosin, 6-( $\gamma\gamma$ -Dimethylallyl)- 7-Methoxycoumarin .. ..	342
		A New Source of "l-Quercitol" (Viburnitol)	514

	PAGE		PAGE
Fensham, P. J.—		Hughes, G. K.—	
Self-Diffusion in Tin Crystals	91	<i>See</i> Anet, E. F. L. J., and	
Self-Diffusion in Polycrystal-		Ritchie, E. . . . .	336, 635
line Tin . . . . .	105	<i>See</i> Anet, E. F. L. J., <i>et al.</i> . .	330
Foster, D. H., Schwerin, Gerda,		<i>See</i> Anet, F. A. L., and	
and Cohen, W. E.—		Ritchie, E. . . . .	346
The Existence of a Uronic		<i>See</i> Ewing, Jean, and Ritchie,	
Acid Ester in Young Wood		E. . . . .	342, 514
of <i>Eucalyptus regnans</i> . .	504	<i>See</i> Gore, P. H. . . . .	136
Galactic Radiation at Radio		Hydrocarbons, The Iodination	
Frequencies . . . . .	19, 234, 251	of Aromatic . . . . .	587
Galactic Radio - Noise at		Indium, A.C. Polarography of . .	567
200 Mc/s., Survey of . . . .	224	Indium Monobromide, The	
Galactic Structure . . . . .	251	Crystal Structure of . . . . .	581
Gore, P. H., and Hughes, G. K.		Iodination of Aromatic Com-	
Studies of the Wallach Trans-		pounds and Hydrocarbons . .	587
formation. I. The Products		Ionization in Linear Polyenes . .	109
of the Reaction . . . . .	136	Jaeger, J. C., and Westfold,	
Guest, P. G.—		K. C.—	
Estimation of the Error at a		Equivalent Path and Absorp-	
Point on a Least-Squares		tion for Electromagnetic	
Curve . . . . .	173	Radiation in the Solar	
Estimation of the Errors of		Corona . . . . .	376
the Least-Squares Poly-		Jenkin, S. R.— <i>See</i> Angyal, S. J.	461
nomial Coefficients . . . .	364	Jurd, L.—	
Gum, C. S.— <i>See</i> Allen, C. W. . .	224	The Iodination of Aromatic	
Gutman, F.— <i>See</i> Breyer, B.,		Compounds. IV. The	
and Hacobian, S. . . . .	558, 567	Iodination of Aromatic	
Hacobian, S.— <i>See</i> Breyer, B.,		Hydrocarbons and Nuclear-	
and Gutman, F. . . . .	558, 567	Substituted Phenols . . .	587
Hancox, N. C.—		Kannuluik, W. G., and Donald,	
The Reaction between		Heather B.—	
Acetone and Ammonia.		The Pressure Dependence of	
II. Isomeric Oximino-		the Thermal Conductivity	
ketones related to Di-		of Polyatomic Gases at 0 °C.	417
acetoneamine . . . . .	450	Labrum, N. R.— <i>See</i> Minnett,	
Heat Conductivity in Gases . .	417	H. C. . . . .	60
Heat Transfer in the Air near		Lahey, F. N., Lamberton, J. A.,	
the Ground, Radiative . . .	274	and Price, J. R.—	
Ho, G.— <i>See</i> Bland, D. E., and		Alkaloids of the Australian	
Cohen, W. E. . . . .	642	Rutaceae. The Structure	
Hope, R. J.— <i>See</i> Lipson, M. . .	324	and Reactions of Acrony-	
Hopper, V. D.—		cidine . . . . .	155
Nuclear Emulsion Technique	512	Lahey, F. N.— <i>See</i> Brown, R. D.	
Hughes, G. K., Neill, K. G.,		593, 615	
and Ritchie, E.—		Lamberton, J. A.— <i>See</i> Lahey,	
The Synthesis of Melicopicine		F. N., and Price, J. R. . . .	155
and Some Trimethoxy-10-		Latitude Effect of Cosmic Rays	
Methylacridones . . . . .	497	at Sea-Level, Remarks on the	183
		Lead, A.C. Polarography of . .	567



	PAGE		PAGE
Least-Squares Curve, Estimation of the Error at a Point on a . . . . .	173	<i>iso</i> Pelletierine and Methyl <i>iso</i> -Pelletierine, A Synthesis of . .	336
Least-Squares Polynomial Coefficients, Estimation of the Errors of the . . . .	364	Phenols, The Iodination of Nuclear-substituted . . . .	587
Lipson, M., and Hope, R. J.—The Synthesis of Polymers in Reduced Wool . . . .	324	4-Piperidone, Some Bridged Derivatives of . . . .	330
Liquid-Vapour Equilibria . .	306	<i>Pleogyne cunninghamii</i> , The Alkaloids of . . . .	346
McCready, L. L.— <i>See</i> Wild, J. P. . . . .	387	Polarization of Thermal "Solar Noise" . . . . .	265
"Marfanil" and its <i>o</i> - and <i>m</i> -Isomers . . . . .	461	Polarography, A.C., of Cadmium, Zinc, Lead, Thallium, Indium, Bismuth, and Antimony . . . . .	567
Marmion, Diana.— <i>See</i> Anet, E. F. L. J., <i>et al.</i> . . . .	330	Polarography with Alternating Currents . . . . .	558, 567
Melicopicine . . . . .	497	Polyenes, Simple Potential Energy Functions, Absorption Spectra, and Ionization in Linear . . . . .	109
Mellor, D. P.— <i>See</i> Stephenson, N. C. . . . .	581	Polymers in Reduced Wool, The Synthesis of . . . . .	324
Methoxycoumarin, 6- ( $\gamma$ -Dimethylallyl)-7-, (Suberosin) . .	342	Precipitation, Observation of, with an Airborne Radar . .	214
Methyl <i>iso</i> Pelletierine and <i>iso</i> -Pelletierine, A Synthesis of . .	336	Price, J. R.— <i>See</i> Lahey, F. N., and Lamberton, J. A. . .	155
Minnett, H. C., and Labrum, N. R.—Solar Radiation at a Wavelength of 3.18 Centimetres . .	60	Priestley, C. H. B.—On the Dynamics of the General Atmospheric Circulation . . . . .	1
Murray, K. E.—A Study of the Oxidation of Ethylene to Ethylene Oxide on a Silver Catalyst . .	433	Proton and Neutron, Interactions between . . . .	519
Naphthoquinones related to Dunnione . . . . .	466, 481, 487	" <i>l</i> -Quercitol" (Viburnitol), A New Source of . . . .	514
Neill, K. G.— <i>See</i> Hughes, G. K., and Ritchie, E. . . . .	497	4-Quinolone, Compounds related to . . . . .	615
Neutron and Proton, Interactions between . . . . .	519	Radar, Observation of Precipitation with an Airborne . .	214
Nicholas, J. F.— <i>See</i> Clarendon, L. M. . . . .	284	Radiation at Metre and Centimetre Wavelengths . . . .	60, 387, 399, 541
Noise (Radio-) at 200 Mc/s., Survey of Galactic . . . .	224	Radiation from the Quiet Sun, Radio-Frequency . . . .	34
Nuclear Emulsion Technique . .	512	Radiation, Galactic, at Radio Frequencies . . . . .	19, 234, 251
Nuclear-substituted Phenols, The Iodination of . . . .	587	Radiation in the Solar Corona, Electromagnetic . . . .	376
Optics, Geometrical, of Concave Mirrors and Combinations of Mirrors . . . . .	532	Rain by Coalescence, The Formation of . . . . .	193
Oximinoketones related to Diacetoneamine . . . .	450		

	PAGE		PAGE
Rathgeber, H. D.—		Suberosin, 6-( $\gamma\gamma$ -Dimethylallyl)-	
Remarks on the Latitude		7-Methoxycoumarin ..	342
Effect of Cosmic Rays at		Sulphonamides ..	461
Sea-Level ..	183	Sun, Quiet, Radio-Frequency	
Ritchie, E.—		Radiation from the ..	34
<i>See</i> Anet, E. F. L. J., <i>et al.</i> ..	330	Sun's General Magnetic Field,	
<i>See</i> Anet, E. F. L. J., and		Determination of the ..	265
Hughes, G. K. ..	336, 635	Swan, P.—	
<i>See</i> Anet, F. A. L., and		Interactions between Neutron	
Hughes, G. K. ..	346	and Proton ..	519
<i>See</i> Ewing, Jean, and Hughes,		Thallium, A.C. Polarography of	567
G. K. ..	342, 514	Thompson, Adrienne R.—	
<i>See</i> Hughes, G. K., and Neill,		A Colorimetric Method for the	
K. G. ..	497	Determination of Esters ..	128
Rutaceae, Alkaloids of the		Tin, Self-Diffusion in Poly-	
Australian ..	155	crystalline ..	105
Schwerin, Gerda.— <i>See</i> Foster,		Tin Crystals, Self-Diffusion in ..	91
D. H., and Cohen, W. E. ..	504	Trimethoxy-10-Methylacridones	497
Segal, W.— <i>See</i> Cooke, R. G. ..	628	Uronic Acid Ester in Young	
Shearer, J.—		Wood, The Existence of a ..	504
Geometrical Optics of Con-		Ursolic Acid from <i>Anthotroche</i>	
cave Mirrors and of Com-		Species ..	516
binations of Mirrors ..	532	Viburnitol ("l-Quercitol"), A	
Silver-Magnesium Alloys, The		New Source of ..	514
Superstructure in the $\alpha$ Phase	284	Wallach Transformation,	
Slee, O. B.— <i>See</i> Stanley, G. J. ..	234	Studies of the. I ..	136
Smerd, S. F.—		Wavelength of 3.18 Centi-	
Radio-Frequency Radiation		metres, Solar Radiation at a	60
from the Quiet Sun ..	34	Wavelengths, Metre, The	
The Polarization of Thermal		Spectrum of High-Intensity	
" Solar Noise " and a De-		Solar Radiation at ..	399, 541
termination of the Sun's		Westfold, K. C.—	
General Magnetic Field ..	265	<i>See</i> Bolton, J. G. ..	19, 251
Smith, E. J.—		<i>See</i> Jaeger, J. C. ..	376
Observation of Precipitation		White, D. E.— <i>See</i> Bottomley,	
with an Airborne Radar ..	214	W. ..	516
" Solar Noise ", Polarization of		Wild, J. P.—	
Thermal ..	265	Observations of the Spectrum	
Solar Radiation at a Wave-		of High-Intensity Solar	
length of 3.18 Centimetres ..	60	Radiation at Metre Wave-	
Somers, T. C.— <i>See</i> Cooke, R. G.		lengths.	
	466, 487	II. Outbursts ..	399
Sparteine and Some Related		III. Isolated Outbursts ..	541
Substances, A Synthesis of ..	635	Wild, J. P., and McCreedy, L. L.	
Stanley, G. J., and Slee, O. B.—		Observations of the Spectrum	
Galactic Radiation at Radio		of High-Intensity Solar	
Frequencies. II. The Dis-		Radiation at Metre Wave-	
crete Sources ..	234	lengths. I. The Apparatus	
Stephenson, N. C., and Mellor,		and Spectral Types of Solar	
D. P.—		Burst Observed ..	387
The Crystal Structure of			
Indium Monobromide ..	581		

	PAGE		PAGE
Williams, E. J.—		<i>Xanthoxylum</i> Species, The	
Experimental Designs		Chemical Constituents of	
Balanced for Pairs of		Australian .. .. .	342
Residual Effects .. ..	351		
Winfield, M. E.—		Yabsley, D. E.—	
The Catalytic Dehydration of		Atmospheric Noise Levels at	
2,3 - Butanediol to		Radio Frequencies near	
Butadiene. II. Adsorption		Darwin, Australia .. ..	409
Equilibria .. .. .	290	Zinc, A.C. Polarography of ..	567





# DATE DUE

RES. 2476

AUG 20 1983

RETD AUG 19 1983

UIC JUN 01 2007

UIC Rec'd MAY 21 2007

GAYLORD

PRINTED IN U.S.A.



3 8198 304 260 621

UNIVERSITY OF ILLINOIS AT CHICAGO

1982  
COPY 1



